AD_____

Award Number: DAMD17-96-1-6084

TITLE: Metastatic Regulation of Differntial Splicing of CD44

PRINCIPAL INVESTIGATOR: Susan M. Berget, Ph.D.

CONTRACTING ORGANIZATION: Baylor College of Medicine
Houston, Texas 77030-3498

REPORT DATE: August 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20000515 092

DTIC QUALITY INSPECTED 4

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1999	Annual (15 Jul 98 - 14 Jul 99)

4. TITLE AND SUBTITLE Metastatic Regulation of Differntial Splicing of CD44

5. FUNDING NUMBERS
DAMD17-96-1-6084

6. AUTHOR(S)
Susan M. Berget, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Baylor College of Medicine
Houston, Texas 77030-3498

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200*

Alternative splicing of pre-mRNAs has the capacity to create many different gene products from a single transcription unit. It has recently become apparent that alternative processing isaltered during cancer. Our project studies the alternative processing of a gene, CD44, that is both massively alternatively processed and in which alterations in alternative processing have been linked to cancer. In the past progress period we have concentrated on identifying and characterizing the factors that bind to two of the alternative exons in the CD44 gene and their expression patterns in a murine model of mammary tumorigenesis. We have discovered that three splicing factors -- human Tra2α, SRp20, and SRp75 affect the inclusion of CD44 alternative exons 4 and 5 *in vivo*. In contrast other known splicing enhancer proteins such as SRp55, ASF/SF2, 9G8, and SC35 have no ability to enhance CD44 splicing. Tra2α is not expressed in mature gland from virgin female mice but is induced in both pregnancy and preneoplasias. High levels of Tra2a were detected in mammary tumors and their metastases, tissues in which the inclusion of CD44 variable exons 4 and 5 increase, suggesting a potential role for Tra2α in the increase of CD44 variable splicing observed during cancer.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer, Splicing Regulation, CD44	39
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

SMB Where copyrighted material is quoted, permission has been obtained to use such material.

SMB Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

SMB In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

SMB In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

SMB In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

SMB In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Susan M. Berget Aug 11, 1999

PI - Signature Date

(4) TABLE OF CONTENTS

(5) INTRODUCTION

Alternative splicing has the capacity to generate a number of RNA isoforms from a single gene. The cell adhesion molecule CD44 can theoretically be alternatively spliced to form more than one thousand different mRNAs (8, 10, 21). The CD44 alternative exons, all in the same translational reading frame, reside within a single block in the middle of the gene (Fig. 1A). Interest in CD44 splicing peaked when it was observed that inclusion of one alternative exon was causative for metastasis in a rat model of colon cancer (9). CD44, is a widely expressed cell adhesion molecule and trans-membrane glycoprotein that mediates a variety of cell-cell and cell matrix interactions (6, 16, 22). A number of CD44 isoforms arising from alternative splicing are expressed in normal tissues (8, 10, 21). Recently it has been shown that alternative splicing of CD44 responds to signal transduction. Treatment of T-lymphoma cells that normally express a form of CD44 lacking variable exons with phorbol esters, concanavalin A, or c-ras induces inclusion of variable exons (13).

We and others have recently observed that CD44 alternative splicing increases in cancer cells (2, 5, 23). There are simultaneous alterations in expression of certain alternative splicing factors, including the well-studied serine-arginine-rich (SR) proteins (7, 23). SR proteins are known to bind to splicing enhancer sequences located within exons and thereby stimulate the binding of constitutive splicing factors such as U2 snRNP-associated factor (U2AF) and U1 snRNPs (1, 17, 20). The transformer 2 protein (Tra2) is one such SR protein (11, 14, 25). Discovered in Drosophila melanogaster, Tra2

regulates alternative splicing in the doublesex gene by binding to CA-rich enhancer sequences (15) within the alternative exon in conjunction with another SR protein, tra.

Recently the human orthologs of Tra2 have been cloned and characterized (3, 24). Two very similar genes Tra2α and Tra2β exist in humans. Both have the capacity to bind to purine-rich exon enhancer sequences although a natural target gene for either human isoform of Tra2 is unknown. Examination of the sequences of CD44 variable exons revealed a number of exon sequences that resembled the CA-rich exon enhancer from the doublesex gene (Fig. 1B-D), suggesting a potential role for human Tra2 in the splicing of CD44. Here we used a mini-gene containing CD44 variable exons v4 and v5 to study the ability of SR proteins, including Tra2, to influence CD44 splicing. We show that Tra2α enhances exon inclusion of both variable exons 4 and 5 in HeLa cells. In contrast, other SR proteins, including SRp55, ASF/SF2, and SC35 had no effect on the splicing of these two exons. One other SR protein, SRp75, also had an effect on splicing of these exons but produced a different phenotype in which only one of the two exons was included in the majority of mRNAs. And finally, SRp20, also stimulated splicing of v4 and v5, although to a lesser extent that Tra2.

In previous studies of the alteration of SR protein expression in tumors, the expression of Tra2 had not been examined. Using a mouse model of mammary tumorigenesis, we followed the expression of Tra2 in both normal development and neoplasia. Tra2 was induced during pregnancy and in preneoplasias. Levels were even higher in neoplasias and their metastases. CD44 alternative splicing, especially of variable exons 4 and 5, paralleled the expression pattern of Tra2, suggesting that changes in SR protein levels during tumorigenesis may contribute to CD44 alternative splicing.

(6) BODY

Experimental Methods, Assumptions, and Procedures

Plasmids and transfections. The reporter CD44 mini-gene used throughout this study (diagrammed in Figure 2B) was created by inserting sequences from the human CD44 gene into the first exon of an engineered β-globin *in vivo* expression plasmid, Dup 33 (4) obtained from R. Kole, University of North Carolina, Chapel Hill, NC. The CD44 sequence spanned a sequence from 792 nucleotides before variable exon 4 to 515 nucleotides after variable exon 5 and included all of the natural exon between exons 4 and 5. The SR protein expression plasmids were obtained from W. Mattox, M.D. Anderson Cancer Center, Houston TX (*D. Melanogaster* tra, human Tra2α and human Tra2β); J. Stevenin, Institut de Genetique et de Biologie Moleculaire et Cellulaire (9G8); J. Manley, Columbia Univ., New York, N.Y (ASF/SF2); and M. Roth, Fred Hutchinson Cancer Research Center, Seattle, WA (SRp75, SRp55, SRp20, SC35).

All transfections used HeLa cells and were done in the presence of lipofectAMINE™ (Gibco/BRL) according to the manufacturer's instructions. Total cell RNA was isolated 48 hours post transfection using TRIzol™ (Gibco/BRL) following manufacturer's instructions. Splicing patterns were determined by RT/PCR analysis using 5' end-radiolabeled primers specific for β-globin sequences (5' AGACACCATGCATGGTGCACC and 3' CCTGATCAGCGAGCTCTAG). These primers amplified no RNA from untransfected HeLa cells. Amplification conditions were 1 min. at 94°C, 1 min. at 58°C and 1.5 min at 72°C for 25 cycles. Product DNA was

denatured and displayed on a 6% urea gel. RNA products were quantified in the Phosphoimager. Plasmid pBR322 digested with Hpa II markers is used in all displayed gels for product size determination. Identified amplification products resulting from the inclusion of one or two variable CD44 exons were sequenced to verify identity.

Tissues. We used an *in vivo* mouse model of mammary development and tumorigenesis to study normal tissue, pregnant and lactating glands, preneoplastic lesions, primary adenocarcinomas and organ metastases to the liver and lung (12, 18). The TM preneoplastic outgrowth lines developed after transplantation of established mouse mammary epithelial cell lines (MMEL) into cleared mammary fat pads of 3 week old syngenic BALB/cMed mice. The serially-transplanted outgrowths were removed either as preneoplasias at 8-12 weeks after transplantation or as tumors 5-7 months after transplantation. After removal, the preneoplastic outgrowths, primary adenocarcinomas, and metastases were frozen at -80 °C for further analysis.

Epithelial cell isolation from mature virgin mammary tissue. Isolated mammary fat pads were minced and incubated in DMEM medium with 2 mg/ml collagenase A (Boehringer Mannheim, Germany) and 100 U/ml hyaluronidase (Boehringer Mannheim, Germany) at 37 °C for 3 hours and slowly swirled. Afterwards the solution was centrifuged for 5 min at 1000 rpm. The supernatant, containing fat and single cells, was discarded and the pellet, containing mammary epithelial cells, was washed in PBS/5% FCS and stored at -80 °C for further analysis.

Analysis of tissue RNA. Tissues were minced on dry ice in before they were treated with a tissue homogenizer (Polytron, Littau, Switzerland) in TRIzol solution (Gibco-BRL, Gaithersburg, MD) with 1ml per 100 mg of tissue for 3 times 10 sec. After 5 minutes of incubation at 26°C, 0.2 ml of chloroform per 1 ml of TRIzol were added and the solution was shaken vigorously, followed by incubation at 26°C for 5 minutes. After centrifugation for 15 min at 4°C and 12.000 x g the aqueous phase, containing the total RNA, was precipitated with 0.5 ml of isopropyl alcohol per 1 ml of TRIzol for 10 minutes at 26°C and centrifuged for 10 min at 12.000 x g at 4°C. The pellets were rinsed with 75% ethanol (1ml of ethanol per 1 ml of TRIzol reagent used for the initial homogenization), vortexed and centrifuged at 4°C for 5 min at 7.500 x g and finally dissolved in RNAse free water and stored at -80°C for further analysis. Five micrograms of RNA were used for cDNA synthesis using M-MuLV reverse transcriptase (Perkin Elmer, Branchburg, NJ) and oligo-dT primers followed by PCR using specific primers: hnRNP A1: 5' GGTGGTCGTGGAGGTGGTT, 3' CCAAAATCATTGTAGCTTCC; Tra2α: 5' TATGATTACCGGTACAGAAGAAGGT, 3' GGAAACAAGCAACACAACTGG; and Tra2β 5' TGACAATACATTTCCACCACAG, 3' GCATCAGCAATTTTTCTTCC

Results and Discussion

A mini-gene containing CD44 variable exons 4 and 5 expresses mRNA including exons 4 and 5 in the presence of some SR proteins. During a previous study of CD44 alternative splicing we discovered that inclusion of variable exon 5 was a good marker for murine mammary tumorigenesis in that this exon was never included in normal RNA but was included in preneoplasia and neoplasias (23). When exon 5 was included, exon 4

was also included. We therefore, decided to begin an analysis of trans-acting factors that regulate CD44 variable splicing by investigating if known splicing factors could influence the level of inclusion of exons 4 and 5. We created a mini-gene in which exons 4 and 5 and the natural intron sequences surrounding them were inserted into intron 1 of a β-globin mini-gene (Figure 2B). Transfected into HeLa cells, only a small percentage of the resulting mRNA included either exon 4 or 5 (Fig. 2A and C).

A number of cDNA expression plasmids coding for known SR proteins were co-transfected with the reporter mini-gene. Plasmids expressing SC35 (Fig. 2A and C) SRp55 (Fig. 5) or ASF/SF2 (Fig. 5) had no effect on the levels of exon inclusion. The lack of effect upon co-transfection with SRp55 is somewhat surprising given the ability of the *D. melanogaster* homologue of SRp55, dSRp55, to interact and be required for doublesex splicing (19). In contrast, co-expression of Tra2α increased inclusion of both exons 4 and 5 (Fig. 2A and C). Total inclusion rose from 27% to 62% +/- 5% in the presence of Tra2α, suggesting that Tra2α could be an alternative splicing factor acting to increase CD44 exon inclusion.

The CD44 variable exons are weak exons. Therefore, it was possible that the effect of Tra2α on exon inclusion was non-specific. To see if Tra2α could increase the inclusion of any weak exon, we tested the ability of co-transfected Tra2α to increase the expression of a purposefully weakened beta globin exon in the same β-globin expression vector used in Figure 2 (Fig. 3). This exon has been used before to study the ability of accessory enhancer sequences and trans-acting factors to regulate splicing (4). Transfection of this mini-gene results in production of mRNA in which less than 20% of the RNA contains the central exon. Co-transfection with Tra2α had no effect on

inclusion of this exon, suggesting that Tra2α cannot bind and affect all weak exons in this mini-gene context (Fig. 3).

Examination of the sequences of exons 4 and 5 revealed the presence of multiple CA-rich sequences within the exons, especially exon 4 (Fig. 1B). These sequences closely resembled the CA-rich sequences in the Drosophila doublesex gene (Fig. 1D), suggesting direct binding of Tra2α to the CD44 exons. To test this hypothesis we asked if purified recombinant Tra2α could bind to exon 4 sequences. Using splicing extracts from HeLa cells, we were able to show that purified recombinant Tra2α bound to exon 4 sequences (data not shown). Both homologous sequences and the Drosophila doublesex enhancer were effective competitors for this binding, but non-specific sequences were not. Thus, the ability of Tra2α to enhance CD44 exon inclusion in HeLa cells is predictable from its RNA binding preference.

Relationship of Tra2α regulation of CD44 splicing to regulation of doublesex splicing. As a positive control for the ability of Tra2α to affect exon inclusion of an appropriate exon we tested the ability of co-transfected Tra2α to increase recognition of the 3' splice site from the *D. melanogaster* female-specific exon in a transfected HeLa cell. It is known that human Tra2α can replace *D. melanogaster* Tra2 *in vivo (3)*. Our experiment was done using a mammalian expression vector in which the second exon began with the 3' splice site of the D. melanogaster doublesex female-specific exon. Transfected into HeLa cells, this reporter is dependent upon co-transfection with Trafor any recognition of the female 3' splice site. In the presence of tra, exon recognition increased; even greater inclusion was observed in the presence of both Traand Tra2α

(data not shown). Thus, human Tra2α has properties very similar to Drosophila Tra2 and would be expected bind to CA-rich sequences.

The observation that human Tra2α affects CD44 alternative splicing suggests that Tra might also be involved. No vertebrate homologue of Tra, however, has been described. It is unclear if such a protein even exists. To test the possibility, we asked if CD44 alternative splicing in our system would respond to addition of *D. melanogaster* Tra into HeLa cells. In this experiment our CD44 reporter plasmid was co-transfected with expression vectors encoding *D. melanogaster* Tra, human Tra2α, or a mixture of the two (Fig. 4). Co-transfection with Tra had a very modest effect on CD44 alternative splicing. More importantly the ability of Tra2α to stimulate CD44 alternative splicing was unaffected by the simultaneous presence of Tra. This observation suggests that the involvement of Tra2α in CD44 splicing is different than its involvement in doublesex splicing.

One other SR protein, 9G8, has been implicated in Drosophila doublesex splicing (14). This protein has an RNA binding domain quite similar to that in both human SRp20 and *D. melanogaster* rbp1. Interestingly, however, transfection of 9G8 had no effect on alternative splicing in our CD44 mini-gene, suggesting another important differences between the alternative splicing of CD44 and the *D. melanogaster* doublesex pre-mRNA (Fig. 2A).

Other SR protein also influence splicing of CD44 alternative exons 4 and 5. Two other SR proteins also affected the levels of exon inclusion in our test mini-gene – SRp20 and SRp75. Like Tra2α, SRp20 co-transfecftion increased the inclusion of both exons

(Fig. 5). The level of inclusion, however, was less than for Tra2α. The *Drosophila* SR protein most-similar to SRp20, rbp1, has been implicated in the recognition of the Drosophila doublesex exon, either through recognition of the exon enhancer or through recognition of the 3' splice site. Thus, it is probably not surprising that SRp20 also activated exon inclusion in CD44 exons containing CA-rich sequences.

The other SR protein that affected splicing in our system was SRp75 (Fig. 5). Unlike the other effectors, however, SRp75 had less ability to increase the inclusion of both exons. Instead, mRNAs were observed that contained one of the variable exons, not both. Because little is known about the RNA binding preferences of SRp75, it is difficult to interpret the effect of SRp75 in this system.

Expression of Tra2α increases during mammary tumorigenesis. Alternative splicing of CD44 increases during the development of murine mammary cancer (2, 5, 23). Thus, any factor postulated to be involved in the alternative processing of CD44 would be expected to have increased or altered activity in tumor cells. To monitor Tra2 expression we utilized a murine system of mammary cancer that we used previously to document increases inCD44 alternative splicing during mammary tumorigenesis (12, 18). RT/PCR analysis was used to determine the splicing patterns of CD44 in both normal and neoplastic cells. Both Tra2α and Tra2β were monitored and compared to an internal standard of an hnRNP protein whose expression is constitutive in most cells. Neither Tra2α or Tra2β protein was detected in normal tissue from virgin female mice (data not shown). Modest levels of both proteins were observed in lactating or pregnant breast, suggesting induction of both forms of Tra2 during pregnancy (Fig. 6). Similar modest

levels were seen in some but not all preneoplasias tested. A consistent high level of both RNAs was observed in cell lines derived from preneoplasias and all tested tumors and their metastases, indicating considerable induction of both factors during tumorigenesis. Thus, Tra2 expression patterns are consistent with a requirement for Tra2 during the increased splicing of CD44 observed during mouse mammary tumorigenesis.

Discussion. Our results indicate that SR proteins may be involved in enhancement of alternative splicing of CD44, especially of the variable exons v4 and v5 whose inclusion increases during many cancers, including breast cancer. Both exons 4 and 5 contain multiple copies of a CA-rich sequence very similar to the CA-rich sequences observed in the *D. melanogaster* doublesex gene, suggesting that some or all of the factors implicated in alternative splicing of doublesex might operate in the alternative splicing of CD44. Here we show that some, but not all, of the proteins implicated in doublesex splicing are factors that can regulate CD44 alternative splicing *in vivo*. Figure 7 contrasts the factors known to bind to the *D. melanogaster* doublesex enhancer as contrasted to the factors described in this report for CD44 exons 4 and 5.

Similar to the regulation of doublesex, Tra2 and the rbp1-like human protein SRp20 activated exon inclusion. Unlike doublesex Tra, 9G8, and SRp55 had no effect on CD44 splicing. In addition one protein not implicated in any known enhancer-mediated splicing event, SRp75, activated exon inclusion of CD44. Our results underscore the modular nature of splicing enhancers. In addition they suggest how combinatorial binding of multiple splicing factors can result in gene-specific splicing patterns.

Recommendations in relation to the statement of work. During the previous funding period our work was intended to concentrate on identifying trans-acting factors that regulate alternative splicing of CD44. As this report shows we are actively demonstrating a number of factors that regulate this alternative splicing event. In addition to the work described above, we have *in vitro* splicing results that indicate that there is at least one additional factor that we need to characterize. Addition of purified recombinant Tra2α to *in vitro* extracts results in increased binding of a non-SR protein of 50 kDA to the CD44 exon 4 CA-rich sequences. We are presently characterizing this factor. We are also currently constructing a battery of exon mutations to sort out which of the factors mentioned in this report bind to which of the possible exon enhancer sequences within exons 4 and 5.

(7) KEY RESEARCH ACCOMPLISHMENTS

- ✓ Establishment of an *in vivo* transfection system that carries out alternative processing of CD44 variable exons and responds to exogenous factors.
- ✓ Determination that human Tra2α influences CD44 alternative splicing
- ✓ Determination that SRp20 and SRp75 also influence CD44 alternative splicing
- ✓ Determination that purified recombinant Tra2α binds to CD44 exon sequences *in vitro.*
- ✓ Determination that recombinant Tra2α activates the binding of one other non-SR protein to CD44 exon sequences
- ✓ Determination that Tra2α is induced during mouse mammary tumorigenesis.

(8) REPORTABLE OUTCOMES

Manuscripts:

Stickeler, E., F. Kittrell, D. Medina, and S. M. Berget. 1999. Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis. Oncogene **18:**3574-3582.

Abstracts

Stickeler, E., D. Medina, T. Cooper, W. Mattox, and S. M. Berget. 1999. CD44 alternative splicing. Cold Spring Harbor Pre-mRNA Processing Meeting, Cold Spring Harbor New York, August 1999.

Employment

Elmar Stickeler has just accepted a position as an Assistant Professor of Obstetrics and Gynecology at the University of Frieberg, Frieberg, Germany and is applying for his first independent funding.

(9) CONCLUSIONS

Our results indicate that alternative splicing of CD44 is a complicated process involving multiple factors that recognize important sequences within the specific exons and boost splicing. Here we report the first function for human Tra2, an important alternative splicing factor in Drosophila used there for sex determination. It's human function has not been known. In addition we report that other SR proteins may interact with Tra2 to promote exon recognition in CD44. Interestingly we observe a different constellation of proteins in CD44 splicing than those reported for D. melanogaster

doublesex splicing. This is the first example of comparison of a set of proteins interacting with similar sequences in two different species. They underscore the ability of multiple splicing factors to bind to a given regulatory sequence and the need to understand the identity of all of the binding proteins to the assessment of function.

(10) REFERNCES

1. **Adams, M. D., D. Z. Rudner, and D. C. Rio.** 1996. Biochemistry and regulation of pre-mRNA splicing. **8:**331-339.
2. **Cannistra, S., G. Abu-Jawdeh, J. Niloff, T. Strobel, L. Swanson, J. Andersen, and C. Ottensmeier.** 1995. CD44 variant expression is a common feature of epithelial ovarian cancer: lack of association with standard prognostic factors. J. Clin. Oncol. **13:**1912-1921.
3. **Dauwalder, B., F. Amaya-Manzanares, and W. Mattox.** 1996. A human homolog of the Drosophila sex determination factor transformer-2 has conserved splicing regulatory functions. Proc. Natl. Acad. Sci. USA **93:**9004-9009.
4. **Dominski, Z., and R. Kole.** 1992. Cooperation of pre-RNA sequence elements in splice site selection. Mol. Cell Biol. **12:**2108-2114.
5. **East, J. A., and I. R. Hart.** 1993. CD44 and its role in tumour progression and metastasis. Eur. J. Cancer **29A(14):**1921-1922.
6. **Fox, S. B., J. Fawcett, D. G. Jackson, I. Collins, K. C. Gatter, A. L. Harris, A. Gearing, and D. L. Simmons.** 1994. Normal human tissues, in addition to some tumors, express multiple different CD44 isoforms. Cancer Res. **54:**4539-4546.

7. **Ghigna, C., M. Moroni, C. Porta, S. Riva, and G. Biamanti.** 1998. Altered expression of heterogeneous nuclear ribonucleoproteins and SR factors in human colon adenocarcinomas. Cancer Res. **58:**5818-5824.

8. **Gunthert, U.** 1993. CD44: a multitude of isoforms with diverse functions. Curr. Top. Microbiol. Immunol. **184:**47-63.

9. **Gunthert, U., M. Hofmann, M. Rudy, S. Reber, M. Zoller, I. Haubmann, S. Matzku, A. Wenzel, H. Ponta, and P. Herrlich.** 1991. A new variant of glycoprotein CD44 confers metastatic potential to rat carcinoma cells. Cell **65:**13-24.

10. **Haynes, B. F., H.-K. Liao, and K. L. Patton.** 1991. The transmembrane hyaluronate receptor CD44: Multiple functions, Multiple forms. Cancer cells **3:**347-350.

11. **Hoshijima, K., K. Inoue, I. Higuchi, H. Sakamoto, and Y. Shimura.** 1991. Control of double-sex alternative splicing by *transformer* and *transformer-2* in Drosophila. Science **252:**833-836.

12. **Kittrell, F. S., C. J. Oborn, and D. Medina.** 1992. Development of mammary preneoplasias in vivo from mouse mammary epithelial cell lines in vitro. Cancer Res. **52::**1924-1932.

13. **Konig, H., H. Ponta, and P. Herrlich.** 1998. Coupling of signal transduction to alternative pre-mRNA splicing by a composite splice regulator. EMBO J. **17:**2904-2913.

14. **Lynch, K. W., and T. Maniatis.** 1996. Assembly of specific SR protein complexes on distinct regulatory elements of the Drosophila doublesex splicing enhancer. Genes Dev. **10:**2089-2101.

15. **Lynch, K. W., and T. Maniatis.** 1995. Synergistic interactions between two distinct elements of a regulated splicing enhancer. Genes Dev. **9:**284-293.

16. **Mackay, C. R., H.-J. Terpe, R. Stauder, W. L. Marston, H. Stark, and U. Gunthert.** 1994. Expression and modulation of CD44 variant isoforms in humans. J. Cell. Biol. **124:**71-82.

17. **Manley, J. L., and R. Tacke.** 1996. SR proteins and splicing control. Genes Dev. **10:**1569-1579.

18. **Medina, D.** 1996. Mammary Tumor Cell Cycle, Differentiation and Metastases, p. 37-69. *In* D. R. a. L. ME (ed.). Kluger Academic Publishers: Norwell.

19. **Peng, X., and S. M. Mount.** 1995. Genetic enhancement of RNA processing defects by a dominant mutation in B52, the Drosophila gene for an SR protein splicing factor. Mol. Cell. Biol. **15:**6273-6282.

20. **Reed, R.** 1996. Initial splice site recognition and pairing during pre-mRNA splicing. Curr. Opin. Genet. Dev. **6:**215-220.

21. **Screaton, G. R., M. V. Bell, D. G. Jackson, F. B. Cornelis, U. Gerth, and J. I. Bell.** 1992. Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. Proc. Natl. Acad. Sci. **89:**12160-12164.

22. **Stamenkovic, I., M. Amiot, J. M. Pesando, and B. Seed.** 1989. A lymphocyte molecule implicated in lymph node homing is a member of the cartilage link protein family. Cell **56: 6:**1057-1062.

23. **Stickeler, E., F. Kittrell, D. Medina, and S. M. Berget.** 1999. Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis. Oncogene **18:**3574-3582.

24. **Tacke, R., M. Tohyama, S. Ogawa, and J. L. Manley.** 1998. Human Tra2 proteins are sequence-specific activators of pre-mRNA splicing. Cell **93:**139-148.

25. **Tian, M., and T. Maniatis.** 1992. Positive control of pre-mRNA splicing in vitro. Science **256:**237-240.

(11) APPENDICES

Figure Legends

Figure 1. Structure and Sequence of the Human CD44 Gene. (A) Diagram of the human CD44 gene with the multiple internal alternative (variable) exons indicated in darker gray. The exons studied in this report are the fourth and fifth alternative exons (variable exons 4 & 5 or v4 & v5). (B & C) Sequences of CD44 variable exons 4 and 5. The exon sequence is boxed. Splice sites are indicated. The repeated CA-rich sequences in each are shown in bold and underlined. A GA-rich repeat known to be an enhancer is shown in italics in variable exon 5. In addition a pyrimidine-rich sequence that is a good candidate binding site for SRp20 is indicated. (D) Comparison of the CA-rich sequences in CD44 variable exons 4 & 5. The CA-rich sequences are complied and compared to

both a derived consensus and the *D. melanogaster* doublesex repeat element known to bind Tra2.

Figure 2. Enhancement of inclusion of CD44 variable exons 4 and 5 by Tra2α. (A) Quantification of RNA resulting from the inclusion of both variable exons 4 and 5 following transfection of HeLa cells with the mini-gene depicted in (B) in the presence of a co-transfected expression plasmid coding for human Tra2α, SC35, or 9G8. Transfections were performed using increasing amounts of expression plasmid and a constant amount of the reporter plasmid (expression plasmid DNA levels used were 0, 1, 2, 3, or 4 μg per 10cm dish of cells). Results shown are the average of three separate transfections. Error bars indicate one standard deviation. (B) Diagram of the reporter plasmid used in this experiment. CD44 variable exons 4 and 5 along with surrounding intron sequences were inserted into the first intron of the human β-globin gene driven by the CMV promoter. Products resulting from exclusion (left) or inclusion (right) of both variable exons are indicated. (C) Results of one of the transfections quantified in (A). Total RNA from a transfection was subjected to RT/PCR amplification using radiolabeled primers specific for β-globin exons 1 and 2. These primers do not amplify any RNA from untransfected cells (data not shown). Amplification products arising from RNA including both one or none of the CD44 exons are indicated to the right of the gel. The extra bands in the Tra2α lanes result from cryptic splicing within the β-globin sequences. Transfections used increasing amounts (0, lanes 1 and 7; 1, lanes 2 and 8; 2, lanes 3 and 9; 3, lanes 4 and 10; or 4, lanes 5 and 11 μg) of indicated expression plasmid. Lane 6 is a size marker.

Figure 3. Tra2α does not increase inclusion of a weak β-globin exon. (A) Quantification of a co-transfection using the reporter plasmid diagrammed in B and an expression vector coding for human Tra2α. (B) The reporter plasmid is Dup33 which includes a weakened internal exon derived from β-globin sequences. The first and third exons in this construct and their flanking splice sites are identical to the mini-gene used in Figure 2. (C) RT/PCR amplification results from the transfection quantified in (A). Product RNAs resulting from inclusion or exclusion of the central exon are indicated.

Figure 4. Drosophila Tra has little effect on CD44 expression. Co-transfections used the CD44 reporter gene diagrammed in Figure 2 and expression plasmids coding for human Tra2α and *D. melanogaster* Tra. (A) Quantification of the amount of inclusion of CD44 variable exons 4 and 5. (B) RT/PCR analysis of RNA products. Products resulting from the inclusion of no, one, or two CD44 exons are indicated. Each transfection used increasing amounts of DNA (0, 1, 2, 3, or 4 μg plasmid). The transfection in the middle used increasing Tra2α in the presence of a constant amount (2 μg) of Tra.

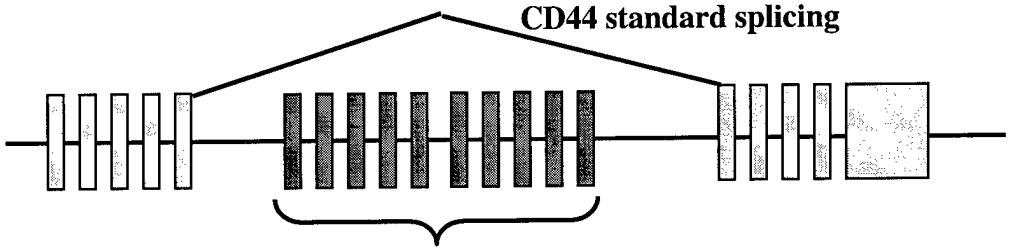
Figure 5. CD44 inclusion is also affected by increasing the level of SRp75 and SRp20. Co-transfections used the CD44 reporter gene diagrammed in Figure 2 and expression plasmids coding for human SRp75, SRp55, SRp20, or ASF/SF2. (A) Quantification of the amount of RNA resulting from inclusion of CD44 variable exons 4 and 5. (B) RT/PCR analysis of RNA products from the transfections with SRp75 (left) and SRp20

(right). Products resulting from the inclusion of no, one, or two CD44 exons are indicated.

Figure 6. Tra2 levels increase during tumorigenesis. RT/PCR was used to quantify the levels of murine Tra2α and Tra2β during tumorigenesis in a mouse model of mammary cancer. Total RNA was extracted from the indicated tissues and cell lines and subjected to RT/PCR amplification using primers specific to Tra2α (A) or Tra2β (B). Amplification of a control gene (hnRNP A1) was used to demonstrate equal RNA isolation from each sample (data not shown).

Figure 7. Model comparing the factors binding to the *D. melanogaster* CA-rich exon enhancer and that found within the CD44 variable exon 4.

A



B

Variable Exon 4

3'

TC**AATCATCGTT**ATCACAG ttt**caaccacaccac**gggcttttg**accaca**

caaaacagaaccaggactggacccagtggaacccaagccattcaaatccg

gaagtgctacttcaga**caaccacaa**ggatgactg GTAATGGGTTCTGCAT

5'

C

Variable Exon 5 3'

TT**TAACCATCATC**ACAG cagatgtagacagaaatggcaccactgcttat

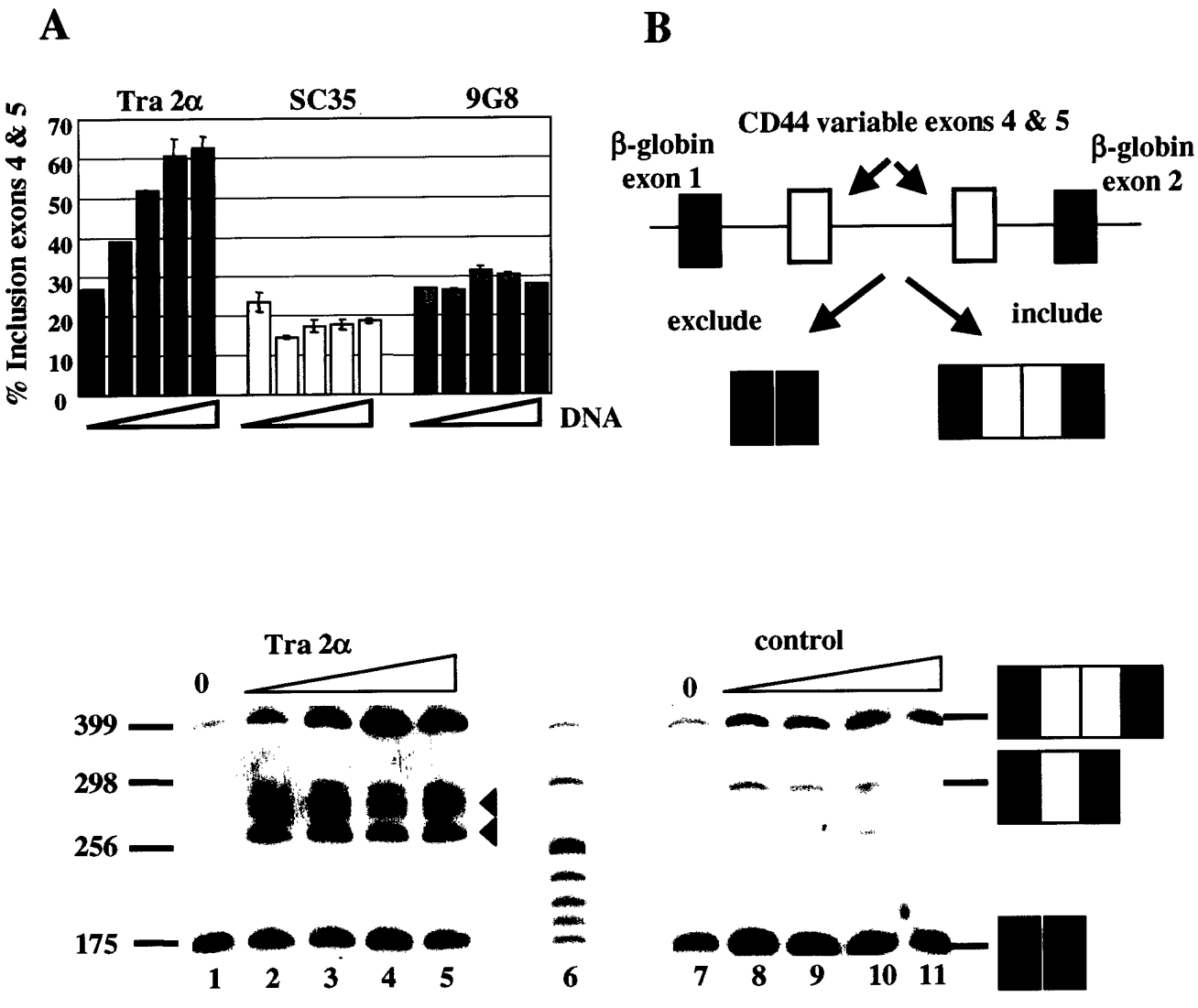
gaaggaaactggaacccagaagcaca*ccctcccctc*attcaccatgagcatca

t*gaggaagaagag*accccacattcta**caagcacaa** GTAAGCAAGATGG

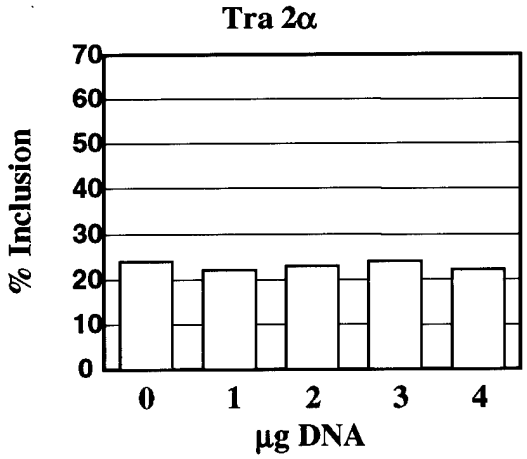
5'

D CA-rich Repeats in CD44 Variable Exons 4 & 5

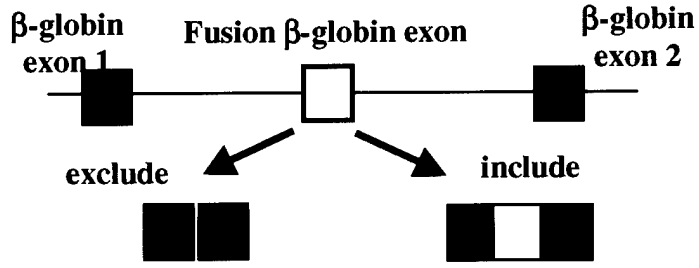
CAATCAACATT	**D. melanogaster doublesex repeat**
CAACCA CAcca	**CD44 exon 4 sequence 1**
gACCA CAcaa	**CD44 exon 4 sequence 2**
CAACCA CAa	**CD44 exon 4 sequence 3**
CAAgCA CAa	**CD44 exon 5 sequence 1**
CAAtCATCgTT	**CD44 exon 4 3' splice site**
tAACCATCATc	**CD44 exon 5 3' splice site**
CAACCA CA	**CD44 Consensus**


A
Tra 2α
SC35
9G8
% Inclusion exons 4 & 5
70
60
50
40
30
20
10
0
DNA
B
CD44 variable exons 4 & 5
β-globin exon 1
β-globin exon 2
exclude
include
Tra 2α
0
399
298
256
175
1
2
3
4
5
6
control
0
7
8
9
10
11

A

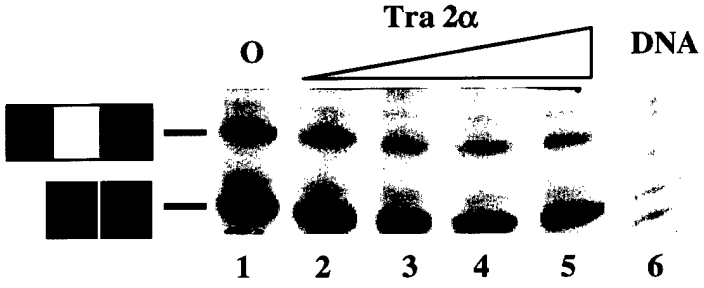


B

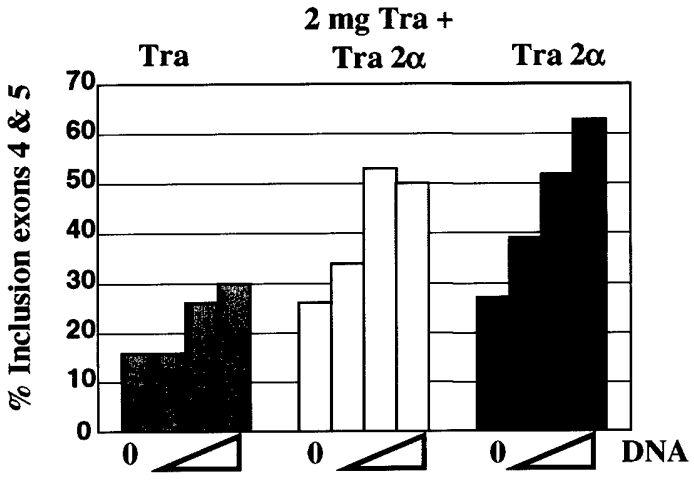


Splicing of Dup 33

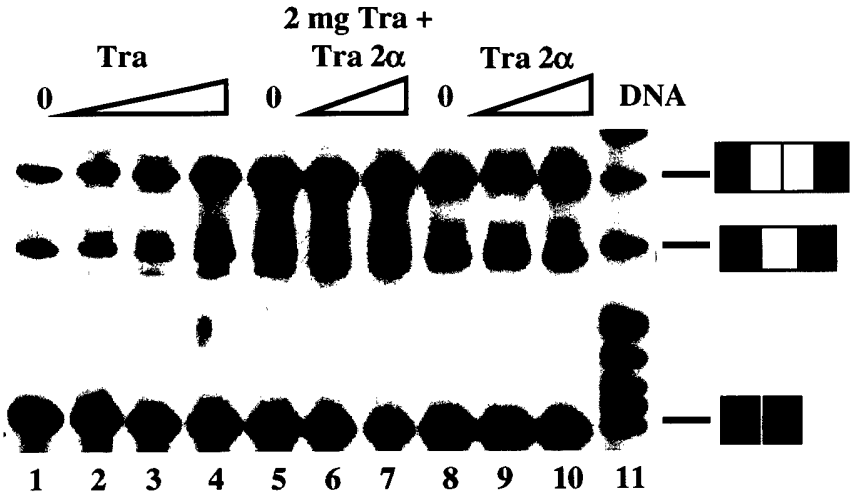
C

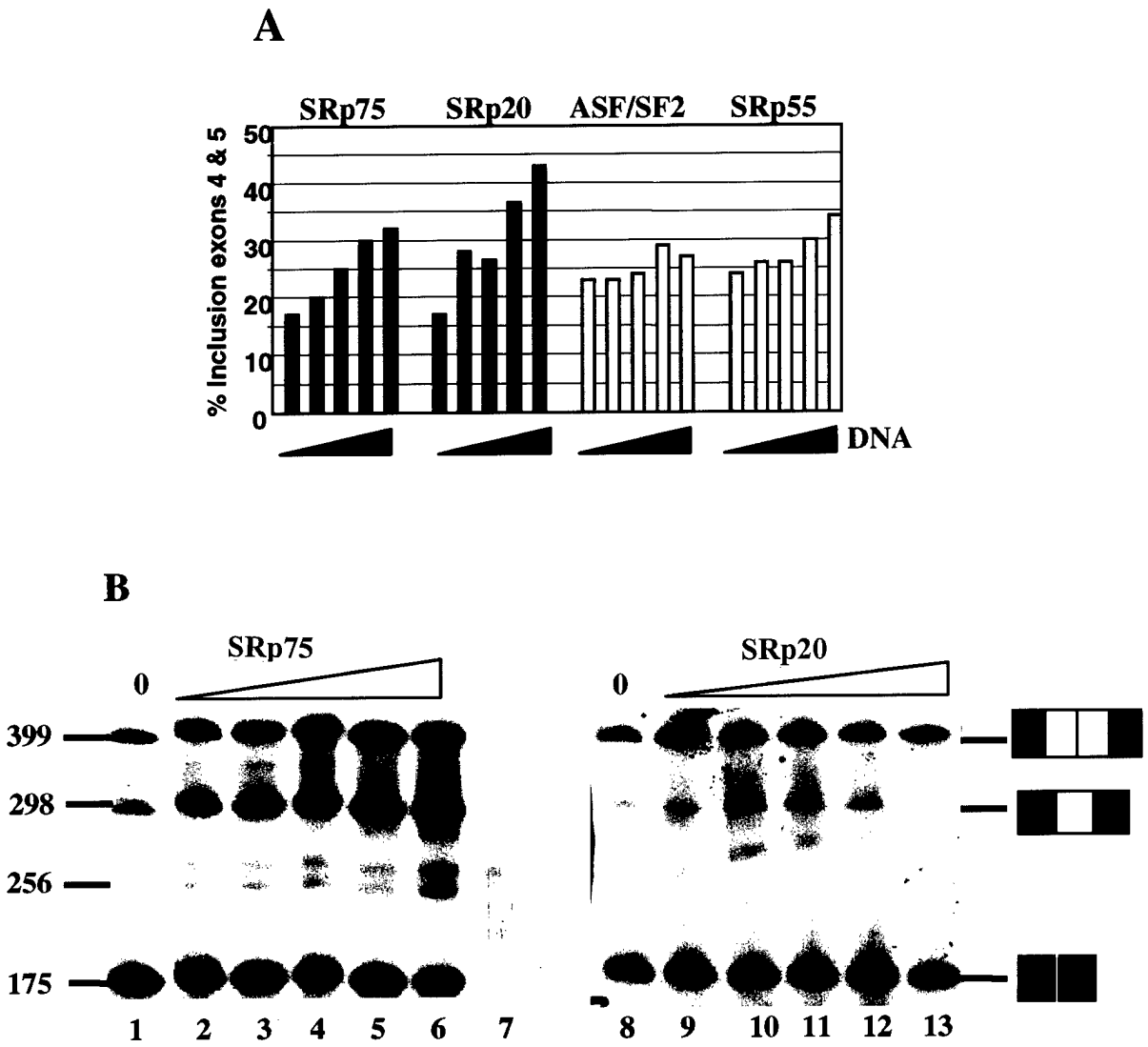


A

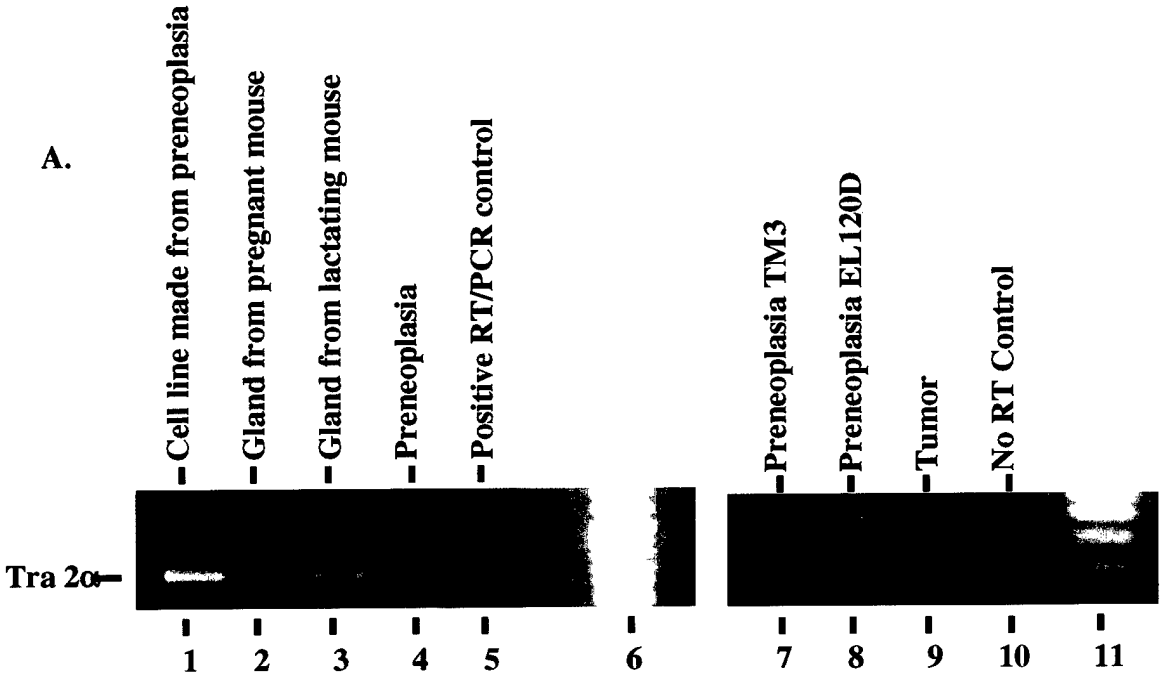

2 mg Tra +
Tra
Tra 2α
Tra 2α
% Inclusion exons 4 & 5
70
60
50
40
30
20
10
0
0
0
0
DNA

B

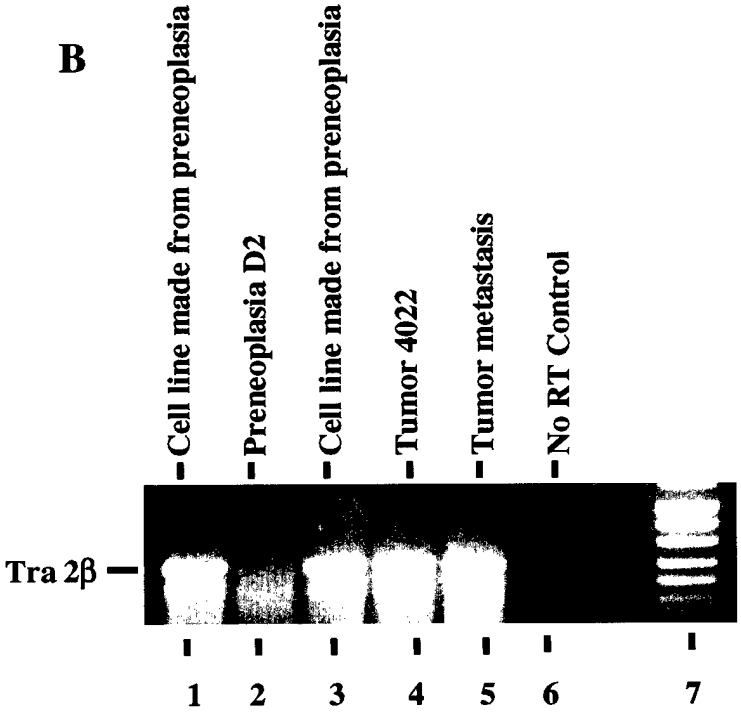

2 mg Tra +
Tra
Tra 2α
Tra 2α
0
0
0
DNA
1
2
3
4
5
6
7
8
9
10
11

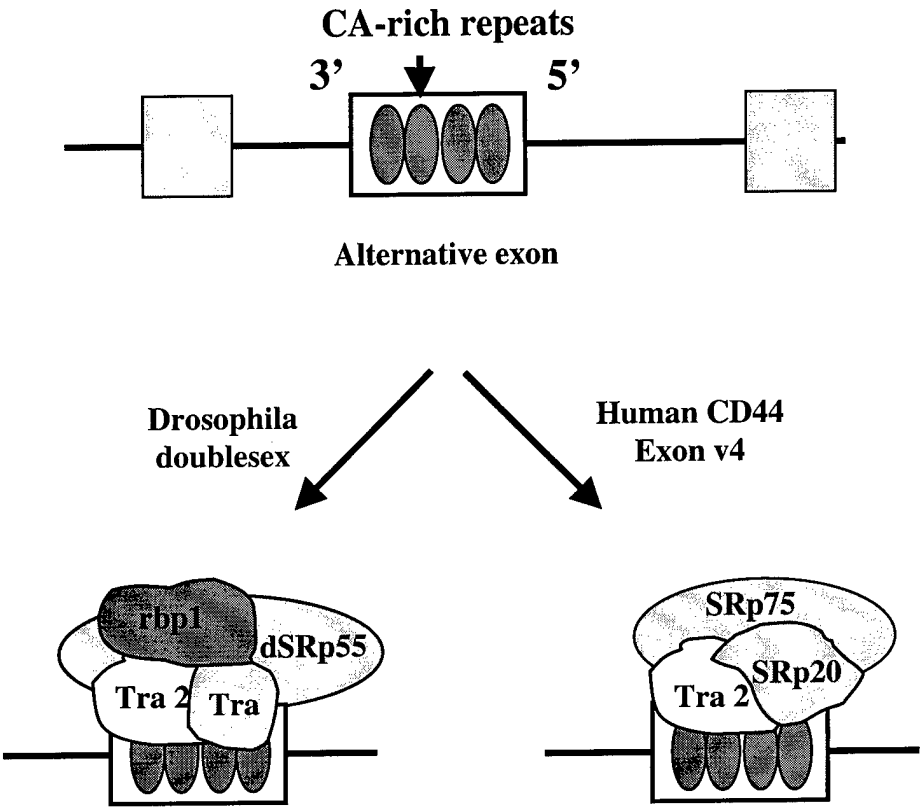

A
SRp75
SRp20
ASF/SF2
SRp55
% Inclusion exons 4 & 5
50
40
30
20
10
0
DNA
B
SRp75
0
SRp20
0
399
298
256
175
1
2
3
4
5
6
7
8
9
10
11
12
13

A.


Cell line made from preneoplasia
Gland from pregnant mouse
Gland from lactating mouse
Preneoplasia
Positive RT/PCR control
Preneoplasia TM3
Preneoplasia EL120D
Tumor
No RT Control
Tra 2α
1
2
3
4
5
6
7
8
9
10
11

B


Cell line made from preneoplasia
Preneoplasia D2
Cell line made from preneoplasia
Tumor 4022
Tumor metastasis
No RT Control
Tra 2β
1
2
3
4
5
6
7



Berget, Figure 7

Oncogene (1999) 18, 3574–3582
© 1999 Stockton Press All rights reserved 0950–9232/99 $12.00
http://www.stockton-press.co.uk/onc

Stage-specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis

Elmar Stickeler[1], Frances Kittrell[2], Daniel Medina[2] and Susan M Berget*,[1]

[1]*Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine, Houston, Texas 77030, USA;* [2]*Department of Cell Biology, Baylor College of Medicine, Houston, Texas 77030, USA*

Using a mouse model of mammary gland development and tumorigenesis we examined changes in both alternative splicing and splicing factors in multiple stages of mammary cancer. The emphasis was on the SR family of splicing factors known to influence alternative splicing in a wide variety of genes, and on alternative splicing of the pre-mRNA encoding CD44, for which alternative splicing has been implicated as important in a number of human cancers, including breast cancer. We observed step-wise increases in expression of individual SR proteins and alternative splicing of CD44 mRNA during mammary gland tumorigenesis. Individual preneoplasias differed as to their expression patterns for SR proteins, often expressing only a sub-set of the family. In contrast, tumors demonstrated a complex pattern of SR expression. Little difference was observed between neoplasias and their metastases. Alternative splicing of CD44 also changed through the disease paradigm such that tumors produced RNA containing a mixture of variable exons, whereas preneoplasias exhibited a more restricted exon inclusion pattern. In contrast, other standard splicing factors changed little in either concentration or splicing pattern in the same cells. These data suggest alterations in relative concentrations of specific splicing factors during early preneoplasia that become more pronounced during tumor formation. Given the ability of SR proteins to affect alternative processing decisions, our results suggest that a number of pre-mRNAs may undergo changes in alternative splicing during the early and intermediate stages of mammary cancer.

Keywords: SR splicing factors; alternative splicing; CD44; mammary tumorigenesis; mouse model

Introduction

Alternate pre-mRNA processing contributes significantly to the developmental regulation of gene expression in humans. It is estimated that as many as 25% of human genes utilize alternative RNA processing to produce subtly or grossly altered gene products (Moore *et al.*, 1993; Rio, 1993; Norton, 1994; Berget, 1995; Kramer, 1996). During the last year, reports have emerged indicating that cancerogenesis induces changes in alternative processing (Lee and Feinberg, 1997; Silberstein *et al.*, 1997; Zhu *et al.*, 1997). Some of the genes targeted for these changes are receptor proteins suggesting that the induced alterations in splicing could have pronounced consequences for cellular behavior. Little, however is known about how RNA processing factors are altered during tumorigenesis so as to affect processing outcomes.

The arginine-serine-rich (SR) proteins (Figure 1a) constitute a family of essential splicing factors (Krainer *et al.*, 1990a; Ge *et al.*, 1991; Zahler *et al.*, 1992) that recognize both splice sites and exonic splicing enhancers, and influence alternative processing decisions when their relative concentrations are altered *in vivo* or *in vitro* (Ge and Manley, 1990; Krainer *et al.*, 1990b; Zahler *et al.*, 1993a; Cáceres *et al.*, 1994; Wang and Manley, 1995). SR proteins have been observed to influence splicing activity via their binding to both splice sites and special splicing accessory sequences known as enhancers (Zahler *et al.*, 1993b; Fu, 1995; Manley and Tacke, 1996; Valcarcel and Green, 1996). Recent studies have indicated substrate-specific binding and activity for individual SR proteins (Fu, 1993; Sun *et al.*, 1993; Zahler *et al.*, 1993a; Cáceres *et al.*, 1994; Wang and Manley, 1995; Chandler *et al.*, 1997; Liu *et al.*, 1998). Furthermore, individual SR proteins have distinct tissue distributions (Zahler *et al.*, 1992; Screaton *et al.*, 1995; Cáceres and Krainer, 1997). These observations have led to suggestions that alterations in the levels of SR proteins could be determinative for alternative splicing during development. This class of proteins, therefore, become attractive candidates for factors whose activity changes during tumorigenesis.

An ideal system in which to study the relationship between processing and the progression of cancer would be one in which both normal and abnormal development could be compared and studied in an isogenic and manipulatable background. Such a system is available in an established *in vivo* mouse model of mammary tumorigenesis (Kittrell *et al.*, 1992; Medina, 1996) that provides access to normal mammary tissues from mature virgin, pregnant and lactating females, as well as an extensively characterized set of preneoplasias and their corresponding adenocarcinomas and metastases. In this model system, preneoplastic outgrowth lines are serially transplanted and maintained for extended periods of time (up to 12 months) in the mammary fat pads of syngenic female BALB/c mice. Each outgrowth line is characterized by a specific rate for development of adenocarcinomas and subsequent metastases (Kittrell *et al.*, 1992; Medina, 1996). The outgrowths are clonal cell populations as determined by oncogene analysis (Cardiff, 1988; Jerry *et al.*, 1993).

Using this system, we characterized SR protein expression during tumorigenesis. At the same time we

*Correspondence: SM Berget
Received 16 July 1998; revised 14 January 1999; accepted 14 January 1999

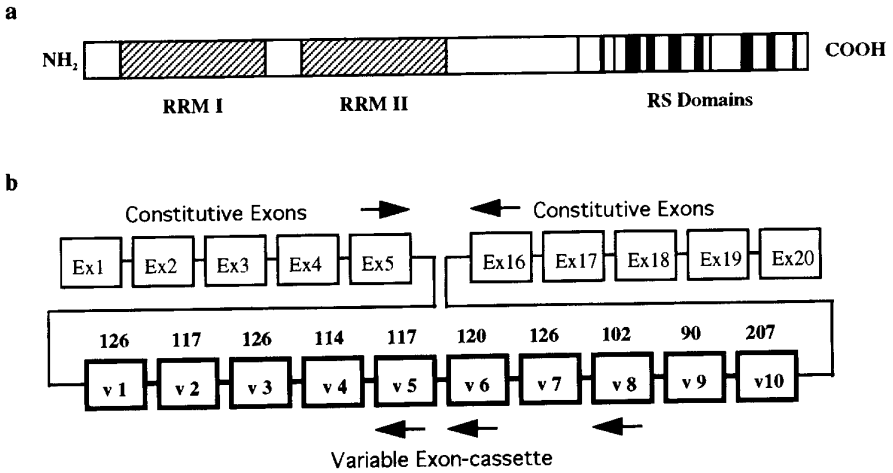


Figure 1 SR proteins and CD44 architecture. (**a**) Schematic of a generic SR protein with two N-terminal RRM (RNA recognition motif) motifs that bind RNA (hatched boxes) and extensive C-terminal RS (arginine-serine) domains (vertical lines). Thickness of vertical lines represents numbers of consecutive RS dipeptides. (**b**) CD44 exon structure. Constitutive exons 1–5 and 16–20 flank a central region encoding at least ten alternatively included (variable) exons (v1–v10). The size of each variable exon is indicated (drawing not to scale). Arrows indicate primers used for RT–PCR analysis

examined the alternative splicing of a pre-mRNA, that coding for CD44, whose splicing had been reported to be determinative for metastasis (Gunthert *et al.*, 1991). CD44, is a widely expressed cell adhesion molecule and trans-membrane glycoprotein that mediates a variety of cell-cell and cell matrix interactions (Fox *et al.*, 1994; Mackay *et al.*, 1994; Stamenkovic *et al.*, 1989). A number of CD44 isoforms are created via alternative pre-mRNA splicing. Within the CD44 gene is an internal cassette of ten alternatively spliced exons, all in the same translational reading frame, and coding for specific extracellular domains of the CD44 protein (Screaton *et al.*, 1992; Figure 1b). Combinations of these variable exons lead to a variety of CD44 isoforms (Screaton *et al.*, 1992; Haynes *et al.*, 1990; Gunthert, 1993). Certain isoforms, especially those including variable exons v5, 6, and 7, have been implicated in the metastasis of several malignancies (East and Hart, 1993; Cannistra *et al.*, 1995; Stickeler *et al.*, 1997; Wielenga *et al.*, 1993) and correlated with survival in human breast cancer (Joensuu *et al.*, 1993; Kaufmann *et al.*, 1995). Recently it has been shown that alternative splicing of CD44 responds to signal transduction. Treatment of T-lymphoma cells that normally express a form of CD44 lacking variable exons, with phorbol esters, concanavalin A, or c-*ras* induces inclusion of variable exons (König *et al.*, 1998). Thus, alternative splicing of CD44 may be involved in malignant transformation of tissues during tumorigenesis, and CD44 joins the collection of genes whose splicing is altered during cancer (Lee and Feinberg, 1997; Zhu *et al.*, 1997; Silberstein *et al.*, 1997).

Recently it has been shown that CD44 alternative splicing may involve inclusion of blocks of alternative exons rather than single isolated exons (Bell *et al.*, 1998). As the metastatic potential of CD44 has been correlated with only a single variable exon, it becomes of interest to see if the pattern of CD44 variable expression during tumorigenesis may have been underestimated by the heavy attention directed to this one exon.

Here we report that step-wise alterations in SR splicing factor levels and CD44 alternative splicing accompanied the transition from normal cells through preneoplasia and mammary adenocarcinoma, but not metastases. Individual preneoplasias exhibited differences in their SR expression pattern and limited CD44 alternative splicing. Alternative splicing of a control gene was not affected during preneoplasia or neoplasia, suggesting gene-specific alterations in RNA processing factors during early tumorigenesis, rather than a general increase in the generic processing machinery. Tumors showed high levels of SR proteins, including proteins never expressed in virgin mammary epithelia and high levels of inclusion of a variety of CD44 alternative exons. Our results suggest that alterations in alternative processing may be important features of early and intermediate stages of mammary cancer.

Results

For this study, a number of mammary preneoplasias and tumors were utilized, including tumors and metastases derived from two of the preneoplasias. The tumorigenic and metastatic properties of two of the preneoplastic outgrowth lines, TM-2L and TM-40, are summarized in Table 1. Both TM-2L and TM-40 produce type B mammary adenocarcinomas by 12 months, at an incidence of 21 or 51%, respectively. The main difference between the two lines is the significant metastatic growth, primarily in lungs and occasionally in liver, of TM-40 (Table 2).

Stepwise alterations in expression of SR proteins during development of mammary cancer

SR proteins are essential splicing factors that participate in early events of substrate recognition by the processing machinery (Fu, 1995; Zahler *et al.*, 1993a,b; Manley and Tacke, 1996; Valcarcel and

Table 1 Tumorigenic characteristics of TM preneoplastic outgrowth lines

Outgrowth line	*Transplant[a] generation*	*Tumors/ transplants (%)*	*TE_{50} (months)[b]*	*% Metastases[c]*
TM2L	21–28	21/98 (21)	9.5	0
TM40	4–13	80/157 (51)	11.0	73

[a]Number of times outgrowth lines were removed from fat pads and retransplanted into cleared fat pads. [b]TE_{50} refers to time for 50% of the transplants to form tumors. [c]Metastases were detected primarily in lung and secondarily in liver in mice 6–8 weeks after surgical removal of the primary tumors (11/15 animals). Metastases are rarely detected in animals bearing primary tumors

Table 2 Preneoplastic origin of adenocarcinomas and organ metastases

Tumor/metastases	*Outgrowth line*
T-5839	TM-2L
T-6788	TM-40
T-7780	TM-40
Liver metastases (T-7780)	TM-40
Lung metastases (T-7262)	TM-40

Green, 1996). In addition they bind to accessory splicing enhancer sequences to facilitate splicing of alternative exons (Ge and Manley, 1990; Cáceres *et al.*, 1994; Wang and Manley, 1995). SR proteins are characterized by an extensive region containing arginine and serine (Figure 1a) which is extensively phosphorylated. Within this phosphorylated domain is the epitope for a monoclonal antibody, mAb104 (Zahler *et al.*, 1993b; Valcarcel and Green, 1996). To examine alterations in splicing factors during development of mammary cancer, we examined the expression of SR proteins in normal and neoplastic mammary tissues by Western blots using mAb104. In HeLa cells, mAb104 recognizes approximately equal amounts of the major SR proteins (SRp20, ASF/SF2, SC35, 9G8, SRp30c, SRp40, SRp55, and SRp75 of molecular weights 20–75 kDa, Zahler *et al.*, 1993b). Mammary epithelial cells from mature virgin glands expressed a sub-set of the major SR proteins, yielding dominant expression of SRp75 and SRp55 (Figure 2a, lane 1). Three of the preneoplasias, TM-2L, TM-4 and TM-10, yielded an SR expression pattern similar to that of mature, virgin mammary gland (Figure 2a, lanes 2, 4 and 5). The other tested preneoplasias, TM-3 and TM-40, in addition, demonstrated detectable levels of SRp40 and SR proteins of the 30 kDa class, suggesting induction of synthesis of these SR proteins during the preneoplastic process (Figure 2a, lanes 3 and 6). Tested adenocarcinomas and their metastases resembled the expression pattern of HeLa cells, characterized by expression of a broad spectrum of SR proteins detected by mAb104 (Figure 2a, lanes 7–9 and data not shown). We have tested a number of independently derived neoplasia for their SR expression profile (data not shown). All demonstrated SR expression patterns identical to those shown in Figure 2a, lanes 7–9, indicating that SR protein expression is induced during tumorigenesis regardless of the expression pattern during pre-neoplasia. Thus, the pattern of SR expression became more complex as neoplasia

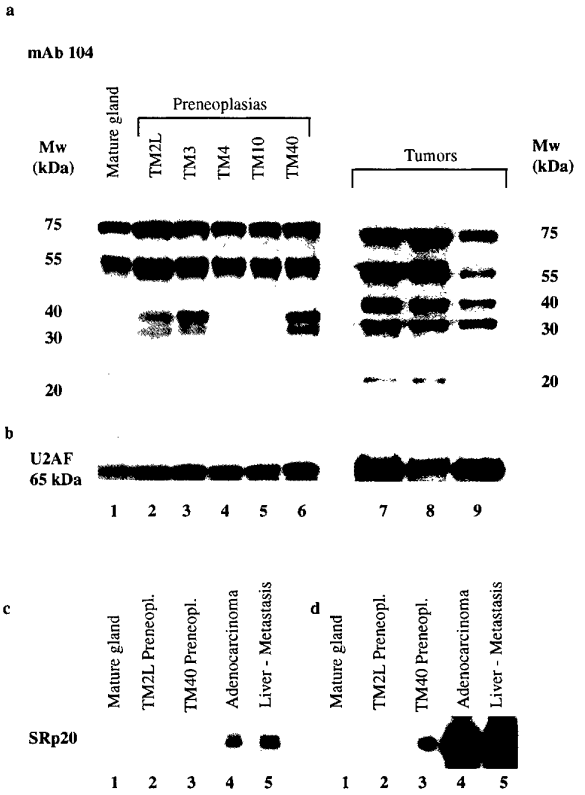


Figure 2 The expression of SR proteins is altered in mammary adenocarcinomas and in some preneoplasias. (**a**) Western blot analysis of SR protein expression in different breast tissues using an antibody specific for a phosphorylated SR epitope (mAb104). Lane 1, mature, virgin gland; lanes 2–6, preneoplasias TM-2L, TM-3, TM-4, TM-10 and TM-40, respectively; lanes 7–9 mammary adenocarcinomas T-5839, T-7780 and T-6129, respectively. Twice as much protein (100 μg) was analysed for normal and preneoplastic tissue (lanes 1–6) than for tumors (50 μg, lanes 7–9). (**b**) Western blot analysis of the same tissue analysed in (**a**) for the presence of the constitutive splicing factor U2AF using a polyclonal rabbit antibody raised against a recombinant fragment of the human 65 kDa subunit missing the SR domain (Zamore *et al.*, 1992). Lanes as in (**a**). (**c**) Western blot analysis of selected tissues from (**a**) using an antibody specific for SRp20. The utilized antibody (Neugebauer and Roth, 1997) is an anti-peptide antibody specific for the region of SRp20 between the RNA binding domain and the SR domain. Lane 1, mature virgin gland; lane 2, preneoplasia TM2L; lane 3, preneoplasia TM40; lane 4, mammary adenocarcinoma T-5839; lane 5, liver metastasis. (**d**) Alterations in levels of mRNAs coding for splicing factor SRp20 during tumorigenesis. Levels of SRp20 revealed by RT/PCR using primers specific for SRp20 (see Materials and methods). Equal amounts (5 μg) of total cell RNA were used for each lane. Lanes as in (**c**). See Figure 3b for RT/PCR analysis of a control RNA in these RNA samples

progressed until a complex array of SR proteins were expressed in tumor cells.

The expression of several control RNA binding factors was monitored over the same disease paradigm. Individual preneoplasias showed identical levels of the 65 kDa subunit of U2AF (Figure 2b), hnRNP A1 (Figure 3b) and hnRNP A2 (Figure 4), as well as small nuclear ribonucleoproteins (snRNPs) (data not shown), indicating that there was not a general increase in the expression of factors involved in pre-mRNA splicing during preneoplasia. Tumors did show a slight increase (estimated to be twofold) for many factors as

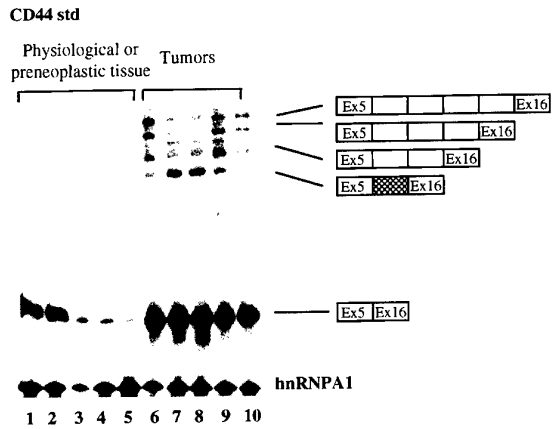


Figure 3 (**a**) Alternative splicing of CD44 changes during mammary development and neoplasia. RT–PCR analysis of CD44 mRNA isoforms in breast tissue using primers specific for constitutive exons 5 and constitutive exon 16. Amplification was for 25 cycles. Boxes symbolize included exons in the amplification bands. Exon structures of amplifications products are indicated. Hatching indicates mixtures of included variable exons. Lane 1, lactating gland; lane 2, pregnant gland; lane 3, mature virgin gland; lane 4, preneoplasia TM-2L; lane 5, preneoplasia TM-40; lanes 6–10, adenocarcinomas T-4031, T-4032, T-5839, T-6129, T-7780, respectively. All lanes used equal amounts (5 μg) of total cell RNA for amplification. (**b**) RT/PCR analysis of a control RNA coding for hnRNP A1 mRNA. Samples are identical to those in (**a**). The form of hnRNP AP shown is that coding for the version of the mRNA lacking exon 7b

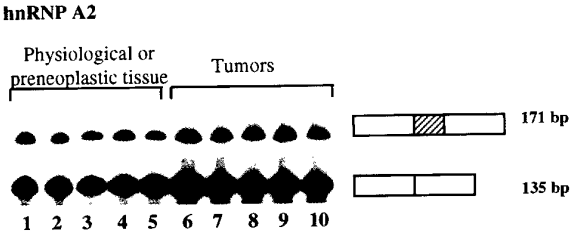


Figure 4 Alternative splicing of mRNA coding for hnRNP A2 does not change through mammary development or neoplasia. Low cycle RT–PCR of hnRNP A2 mRNA in mammary tissues using primers specific for two constitutive exons (white boxes) flanking a 36 nucleotide alternatively included exon (hatched box). Exon skipping produces a 135 nt band (bottom), inclusion in a 171 nt amplification product (top). Lane 1, lactating gland; lane 2, pregnant gland; lane 3, mature virgin gland; lane 4, preneoplasia TM-2L; lane 5, preneoplasia TM-40; lanes 6–10, adenocarcinomas T-4031, T-4032, T-5839, T-6129 and T-7780, respectively. Equal amounts of total cell RNA (5 μg) were analysed

compared to the preneoplasias, including the SR proteins SRp75 and SRp55 (Figure 2a). More interesting, however, was the pronounced induction of expression of the lower molecular weight SR proteins in the 20–40 kDa size range during neoplasia.

One of the SR proteins induced upon tumorigenesis was SRp20. This protein is only weakly detected by mAb104. By Western blotting, this protein was essentially undetectable in preneoplasias. Neoplasias weakly showed production of SRp20. To better examine induction of this factor we used an antibody specific for SRp20. This antibody was raised against an SRp20 peptide from the region of the protein located between the RNA-binding domain and the phosphorylated SR region (Neugebauer and Roth, 1997). As shown in Figure 2c, SRp20 levels detected by this antibody resembled those detected by the mAb104 antibody. No expression of SRp20 protein was detected in virgin or preneoplastic gland. In contrast, SRp20 was expressed in tumor tissue. To rule out the possibility of a false positive resulting from the presence of endothelial cells or infiltrating inflammatory cells in the excised adenocarcinomas, we analysed SR protein levels in cultured mammary epithelial cells established from the tumors in this study. In parallel with the tumours, each cell line demonstrated high levels of the entire family of SR proteins recognized by mAb104 (data not shown).

To provide more sensitive detection of SRp20, we monitored SRp20 mRNA levels using RT–PCR (Figure 2d). To control for non-linear estimations of mRNA levels using RT–PCR, SRp20 mRNA levels were also determined by RNase protection (data not shown). As observed by analysis with the Western blotting, tumorigenesis was accompanied by a major increase in the levels of SRp20 mRNAs, indicating that the increased levels of SRp20 in tumors seen with Western blotting reflected increased mRNA synthesis rather than increased phosphorylation of an existing protein population (Figure 2d).

RT–PCR revealed the presence of a low level of the tested SRp20 mRNAs in the TM-40 preneoplasia that had not been detected by Western blotting, suggesting that low level synthesis of SRp20 occurs in the TM-40 preneoplasia. Even using the power of PCR amplification, however, no SRp20 was detected in preneoplasia TM-2L or in mature virgin tissue. These results indicate that SRp20 is induced during development of mammary carcinogenesis. Furthermore, it indicates a progressive increase in expression as normal cells alter to preneoplasias and then to tumors.

The splicing of CD44 changes during progression of mammary cancer

Alteration in the relative levels of members of the SR proteins recognized by mAb104 is known to affect splicing phenotypes. Thus, the changes in overall amount and, more importantly, in relative amounts of the SR proteins displayed in Figure 2 should be accompanied by changes in alternative processing of target genes. To examine this possibility we looked at the splicing of CD44, a gene with multiple alternative exons. Some of these exons have sequences similar to the purine-rich splicing enhancers known to be the binding sites for the mAB104-reactive SR proteins (Black, 1995; Manley and Tacke, 1996; Reed, 1996; Rio, 1993; Bell *et al*., 1998; Liu *et al*., 1998). Splicing of CD44 was monitored by RT–PCR analysis of RNA from the same tissues used to examine SR proteins. CD44 mRNA is expressed as two basic forms, with or without one or more of the alternative internal exons. Expression of CD44 standard lacking all variable exons (CD44 std) was examined using primers complementary to constitutive exons 5 and 16, which border the central region of the CD44 gene encoding the ten alternative exons (Figure 1b). RT–PCR of CD44

standard mRNA should produce an amplification band of 221 nucleotides; inclusion of one or more variable exons should produce larger bands. Because all of the variable exons are approximately the same size (Figure 1b) inclusion of multiple variable exons produces an amplification pattern with a ladder of bands differing in length by the average variable exon length (approximately 115 nucleotides). This approach permits evaluation of the overall amount of CD44 mRNA present in examined tissues as well as an estimation of the extent of alternative splicing. Amplification revealed that mRNA encoding for CD44 std, lacking the variable exons, was expressed in all tissues examined (Figure 3). Tumor RNA from adenocarcinomas or their metastases demonstrated noticeable alternative processing to reveal a family of amplification bands resulting from the inclusion of multiple variable exons.

To distinguish CD44-specific effects on alternative splicing associated with neoplastic transformation from generic effects on splicing activity associated with an accelerated growth rate, we examined the amounts and spliced isoforms of an RNA that undergoes stochastic alternative splicing in many tissues. The hnRNP A2 gene is processed to produce two mRNA isoforms which alternatively include a 36 nucleotide variable exon. We observed relatively equal levels of total hnRNPA2 mRNA in all tissues examined and approximately equal relative levels of inclusion of the variable exon (Figure 4). This result suggests that the changes in CD44 splicing occurring in mammary neoplasia is due to a change in the sub-set of splicing factors required to recognize the CD44 alternative exons and not a change in the generic splicing machinery.

CD44 metastatic exons are expressed at different levels during normal mammary gland development and tumorigenesis

Low-cycle RT–PCR with primers specific for constitutive exons 5 and 16 produces complicated amplification patterns because of the potential complexity of the family of CD44 mRNA sequences. To better examine variable exons splicing, RNAs from tissue samples were analysed by RT–PCR using exon-specific primers for variable exons 5, 6, 7 or 8 and a consistent primer for constitutive exon 5 (Figure 5a–d), followed by cloning and subsequent sequencing of the PCR products. This approach allowed examination of the inclusion of all alternative exons in the interval v1–v8. Both low and high cycle amplifications were performed. We concentrated on analysis of species containing variable exons v5, 6 or 7, because these exons have been associated with metastasis in other studies (Stickeler *et al.*, 1997; Heider *et al.*, 1995; Sinn *et al.*, 1995). Variable exon v8 was also of interest because it is one of three so-called 'epithelial ' exons, included in CD44 mRNA in a variety of epithelial cells.

Low cycle RT–PCR analysis of RNA from multiple tumors or their metastases using a primer specific for variable exon 5 (v5) indicated considerable production of RNAs containing v5 with or without other upstream variable exons (Figure 5a and data not shown). The inclusion patterns for all of the examined tumors and metastases were very similar, and indicated that most of the observed RNA included multiple variable exons from the v1–v5 region. In contrast, levels of inclusion of v5 were lower in normal dividing cells from pregnant or lactating mammary gland and in the two tested preneoplasias. Using 25–35 cycles of RT–PCR (Figure 5a and b), the major amplified species in these tissues included either v5 alone or v5 and one other upstream variable exon. Cloning and sequencing of the larger amplification band indicated that 4/5 sequenced clones contained v4 in addition to v5. This species was more obvious in pregnant and lactating RNA than in RNA from the preneoplasias (Figure 5a). We were unable to detect any RNA containing v5 in RNA isolated from mature virgin mammary gland (Figure 5a, lane 3; and 5b, lane 2). Because of the abundance of fat tissue in mature virgin gland, we isolated epithelial cells from the extracted virgin mammary tissue. High cycle amplification of this RNA yielded no demonstrable inclusion of exon v5 either (data not shown). Therefore, for v5 splicing, three patterns of inclusion were observed by RT–PCR and sequencing: high level of inclusion of multiple exons along with v5 in tumor cells, low level of inclusion of v5 or v5+4 in preneoplasias, and no inclusion of v5 in normal virgin mammary gland. This pattern suggests step-wise alterations in CD44 alternative splicing during neoplasia. In addition, our data suggest that preneoplasias adopt a splicing pattern similar to that induced in mammary epithelia during normal mammary development and that tumorigenesis is accompanied by induction of a pattern of frequent alternative splicing not found in normal development.

Analysis of alternative splicing of v6 or 7 gave results similar to that discussed above for v5 (Figure 5c and data not shown). Inclusion of both exons occurred at high frequency along with other upstream exons in tumor cell RNA. RNA from preneoplasias and pregnant gland demonstrated simpler RNAs containing fewer variable exons. Sequencing of PCR products indicated that for v6, these RNAs contained only v6. For v7, species containing v7 or v7+4 were observed. In contrast to v5, both v6 and 7 were included in RNA from mature virgin gland epithelial cells (Figure 5c, lane 5; and data not shown). Like v5, these observations suggest that splicing in preneoplasias is similar to splicing in normal proliferative mammary gland and that the conversion to neoplasia is accompanied by pronounced increased in mRNA including these variable exons. Therefore, for this system of mammary neoplasia, inclusion of v6 is not a marker for either tumorigenesis or metastasis. Instead, a better marker is increased inclusion of a number of CD44 variable exons including v6.

A recent report (Bell *et al.*, 1998) indicated that inclusion of blocks of variable CD44 exons was a frequent phenotype in mice. Such multiple inclusion events are certainly consistent with the patterns we observed in the tumors. Preneoplasias and mature virgin mammary epithelial cells, however, demonstrated inclusion of individual exons without inclusion of an immediate neighbor. For example, v6 was included without v5 in mature gland and preneoplasias and v7 was included with v4 but not v6 and 5. If inclusion of multiple exons reflects a mechanism change with respect to inclusion of single exons, then

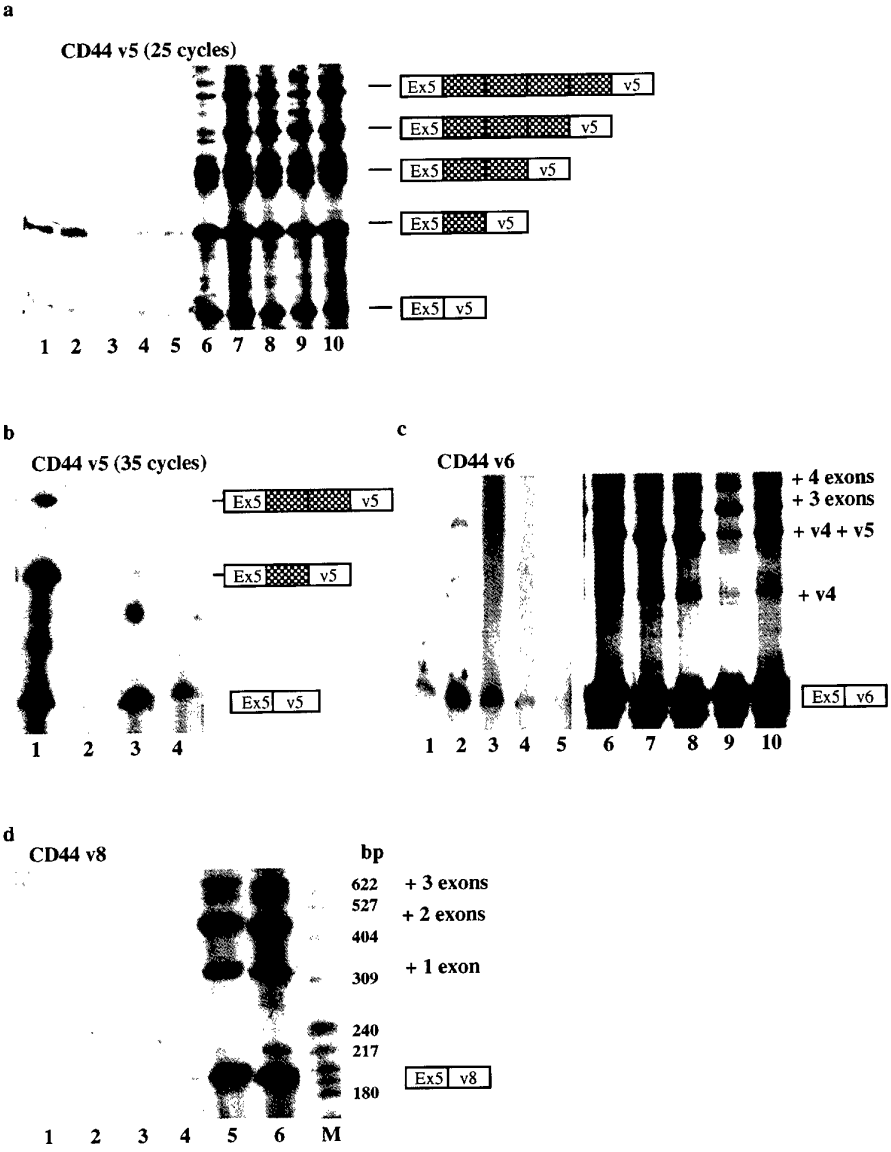


Figure 5 CD44 metastatic exons are expressed in normal development and tumorigenesis. RT–PCR analysis of CD44 mRNA variable isoforms using primers specific for constitutive exons 5 and variable exon v5 (**a** and **b**), variable exon v6 (**c**) or variable exon v8 (**d**). Amplification was for 25 (**a**, **c**, **d**) or 35 (**b**) cycles. Boxes symbolize included exons in the amplification bands. Exon structures of amplifications products are indicated. Hatching indicates mixtures of included variable exons. Majority products are indicated where they existed. (**a**, **c**) Lane 1, lactating gland; lane 2, pregnant gland; lane 3, mature virgin gland; lane 4, preneoplasia TM-2L; lane 5, preneoplasia TM-40; lanes 6–10, adenocarcinomas T-4031, T-4032, T-5839, T-6129 and T-7780, respectively. (**b**) Lane 1, pregnant gland; lane 2, mature virgin gland; lane 3, preneoplasia TM-2L; lane 4, preneoplasia TM-40. (**d**) Lane 1, lactating gland; lane 2, mature virgin gland; lane 3, preneoplasia TM-2L; lane 4, preneoplasia TM-40; lanes 5–6, adenocarcinomas T-4031 and T-4032. All lanes used equal amounts (5 μg) of total cell RNA for amplification

tumorigenesis is accompanied by this mechanism change.

Analysis of RNA for the presence of v8 yielded a surprising result. Like the other variable exons, v8 was included in a number of RNA species in mammary tumor cells (Figure 5d, lanes 5–6). Unlike v5–7, however, no v8 inclusion could be observed in preneoplasias or RNA from developing mammary, despite their epithelial origin (Figure 5d, lanes 3–4). The observation of induction of v8 splicing in tumor RNA indicates that tumorigenesis is accompanied by alterations in RNA processing of CD44 that are never seen in normal development.

Discussion

During the last several years, studies have suggested changes in pre-mRNA splicing in human malignancies (Lee and Feinberg, 1997; Zhu *et al*., 1997; Silberstein *et al*., 1997). Using an *in vivo* model of mammary development and tumorigenesis, we were able to study these questions with a focus on comparing splicing phenotypes and factors during normal mammary development to those occurring in development of mammary cancer. Neoplasia was accompanied by a dramatic increase in expression of some of the SR family of splicing factors recognized by mAb104,

resulting in alteration of the relative abundance of individual SR proteins and an increase in the complexity of expression of this important class of splicing factors. Given the number of genes whose splicing has been observed to be responsive to relative SR protein levels (Ge and Manley, 1990; Krainer *et al.*, 1990b; Cáceres *et al.*, 1994; Wang and Manley, 1995), this observation suggests that pronounced changes in alternative splicing of a number of pre-mRNAs should accompany mammary tumorigenesis.

The utilized model system also permitted examination of SR protein expression in preneoplasias with differing capacity to form tumors or for those tumors to metastasize. The preneoplasias characterized in this study differed in their SR protein expression pattern. Some of the neoplasias had mAb104-reactive SR expression patterns similar to that seen with virgin mammary tissue. The tested preneoplasias therefore appeared to represent different stages in the transition from normal to tumor tissue with respect to SR protein expression. The utilized preneoplasias arose independently and have different tumor and metastasis producing capabilities. SR expression levels did not correlate to tumor or metastatic incidence frequency. Thus, although SR expression changes may be a marker for preneoplasia, they are not predictive for tumor incidence or invasiveness.

Analysis of other constitutive splicing factors indicated that the changes in SR expression we observed were not the result of induction of all splicing during neoplasia. Levels of the constitutive factors U2AF65, two hnRNP proteins, and U snRNPs were only modestly altered during preneoplasia and neoplasia. Therefore, the alterations we observed occurred in a family of factors associated with alternative splicing and the recognition of exon enhancer sequences. This observation suggests that the splicing of multiple target genes regulated by SR proteins would be expected to be altered in neoplasia, without a general effect on exon inclusion levels.

We also used this model to investigate potential changes in alternative splicing in a gene whose splicing changes during cancer have been correlated to metastatic potential in human cancers and which has recently been shown to alter its splicing pattern in response to signal transduction (König *et al.*, 1998). As with SR proteins, striking increases in alternative splicing of CD44 were observed in adenocarcinomas as compared to either preneoplasias or proliferative mammary epithelia from pregnant animals. All of the tumors examined displayed a similar pattern of exon inclusion, in which multiple alternative exons were included in mRNA. Included in this set of exons were exons not normally present in epithelia of pregnant mammary gland or in preneoplasias. The splicing pattern of the adenocarcinomas was similar to that of the lung and liver metastases of these tumors, indicating few changes accompanying metastasis. In contrast to CD44 alternative splicing, the splicing of a control RNA did not change during the transition from preneoplasia to neoplasia. This difference suggests that the pronounced changes in splicing of CD44 observed during development of neoplasia arise from the induction of specialized splicing factors, not an increase in the generic splicing machinery.

Although pronounced changes in CD44 splicing correlated with alterations in the expression level of SR proteins during the tumorigenesis paradigm used in this study, it is not yet clear if the two are related. Transfection of cells that normally skip CD44 variable exons 4 and 5 with cDNAs for SRp55 and ASF/SF2 had little effect on exon inclusion in a co-transfected mini-gene containing these two exons (data not shown). In contrast, transfection with cDNAs for both SC35 and tra2 did alter the percentage of exon inclusion (Stickeler and Berget, manuscript in preparation). Therefore, the splicing of CD44 may be regulated by a sub-set of the SR family of splicing factors.

When we examined CD44 variable exon splicing in epithelia from pregnant mammary gland, we observed low levels of inclusion of some but not all the variable exons. In particular, v6 and 7 were present in a subpopulation of CD44 mRNA from both pregnant tissue and preneoplasias, indicating that these so-called 'metastatic' exons are included in non-cancerous mammary epithelial cells. In contrast, one variable exons, v5, was not expressed in virgin mammary gland, indicating that alterations in CD44 alternative splicing can occur during preneoplasia and resemble those occurring during pregnancy. Other variable exons that are not included in pregnant tissue (such as the 'epithelial' exon v8) are also not expressed during preneoplasia, but are included in tumor RNA. Therefore, our analysis suggests that preneoplasias demonstrate a splicing pattern resembling that occurring during normal breast development. Known to participate in cell-cell interactions and cell-matrix interactions, CD44 isoforms could be involved in the ductal outgrowth occurring during early pregnancy.

Inclusion of three of the CD44 variable exons (v5, 6 and 7) have been postulated to be markers for metastasis potential in several human cancers, including breast cancer. The initial interest in these exons arose from an elegant experiment demonstrating that transfection of a CD44 cDNA including exon v6 into a tumor cell of no metastatic potential increased the rate of metastasis when these cells were placed back into syngenic animals. Such experiments expressed the variant CD44 mRNA at high levels. Our analysis would suggest that the major change in CD44 splicing occurs during development of adenocarcinomas, not their metastases and that it may be the over-expression of variant CD44 mRNAs during the neoplastic state which is important for the metastatic process, rather than the production of any mRNA including v6. This difference may suggest why experiments monitoring inclusion of CD44 variable exons have failed to demonstrate a consistent correlation to disease progression in several clinical studies (Friedrichs *et al.*, 1995; Muller *et al.*, 1997; Tran *et al.*, 1997).

The recent report of a coupling between signal transduction and alternative splicing of CD44 concentrated on variable exon v5. Inclusion of this exon showed a pronounced response to induction by c-*ras*, TPA, or phorbol esters (König *et al.*, 1998). Exon v5 contains internal sequences that resemble known purine-rich exon enhancers that bind SR proteins. These observations suggest that signal transduction cascades can lead to an induction of SR proteins needed for CD44 variable splicing. In light of these findings, our observation that inclusion of v5 is

changed upon preneoplasia supports a possible connection between SR protein levels and CD44 alternative splicing.

Besides CD44 a number of other genes show altered RNA processing during tumorigenesis. A number of cellular receptors and hormone genes undergo alternative processing and this processing appears to alter during cancer. Because the changes in these proteins caused by alternative processing can have a pronounced effect on cellular function, it is important to understand the role of alternative processing in cancer and the mechanisms involved. Determination of splicing factor constellations patterns in preneoplastic lesions of the mammary gland could be very helpful to identify patients in high-risk situations to develop invasive breast cancer and subsequent organ metastasis. Distinct expression patterns of splicing factors could serve as new markers for metastasis in breast cancer, and splicing factors may represent targets for intervention in a subgroup of patients abnormally expressing certain factors.

Materials and methods

Tissues

We used an *in vivo* mouse model of mammary development and tumorigenesis to study normal tissue, pregnant and lactating glands, preneoplastic lesions, eight primary adenocarcinomas and organ metastases to the liver and lung. The TM preneoplastic outgrowth lines developed after transplantation of established mouse mammary epithelial cell lines (MMEL) into cleared mammary fat pads of 3-week-old syngenic BALB/cMed mice (Kittrell *et al.*, 1992). The serially-transplanted outgrowths were removed either as preneoplasias at 8–12 weeks after transplantation or as tumors 5–7 months after transplantation. After removal, the preneoplastic outgrowths, primary adenocarcinomas, and metastases were frozen at −80°C for further analysis. The biological properties of the outgrowth lines are shown in Table 1.

Epithelial cell isolation from mature virgin mammary tissue

Isolated mammary fat pads were minced and incubated in DMEM medium with 2 mg/ml collagenase A (Boehringer Mannheim, Germany) and 100 U/ml hyaluronidase (Boehringer Mannheim, Germany) at 37°C for 3 h and slowly swirled. Afterwards the solution was centrifuged for 5 min at 1000 r.p.m. The supernatant, containing fat and single cells, was discarded and the pellet, containing mammary epithelial cells, was washed in PBS/5% FCS and stored at −80°C for further analysis.

RNA isolation, RT–PCR and sequencing

Tissues were minced on dry ice before they were treated with a tissue homogenizer (Polytron, Littau, Switzerland) in TRIzol solution (Gibco-BRL, Gaithersburg, MD, USA) with 1 ml per 100 mg of tissue for three times 10 s. After 5 min of incubation at 26°C, 0.2 ml of chloroform per 1 ml of TRIzol were added and the solution was shaken vigorously, followed by incubation at 26°C for 5 min. After centrifugation for 15 min at 4°C and 12 000 *g* the aqueous phase, containing the total RNA, was precipitated with 0.5 ml of isopropyl alcohol per 1 ml of TRIzol for 10 min at 26°C and centrifuged for 10 min at 12 000 *g* at 4°C. The pellets were rinsed with 75% ethanol (1 ml of ethanol per 1 ml of TRIzol reagent used for the initial homogenization), vortexed and centrifuged at 4°C for 5 min at 7500 *g* and finally dissolved in RNAse free water and stored at −80°C for further analysis.

Five micrograms of RNA were used for cDNA synthesis using M-MuLV reverse transcriptase (Perking Elmer, Branchburg, NJ, USA) and oligo-dT primers followed by PCR using primers specific for different SR proteins:

SRp20: 5′-GCCGTGTAAGAGTGG,
3′-AAGCTTCCTCCTTCTTGG;
hnRNP A1: 5′-GGTGGTCGTGGAGGTGGTT,
3′-CCAAAATCATTGTAGCTTCC, and
hnRNP A2: 5′-ACAGTCTGTAAGCTTTCCCC,
3′-CTGAAGCGACTGAGTCCGCG.

The standard form of CD44 mRNA (CD44 std) was detected using PCR primers for constitutive exons 5 and 16:

CD44 exon 5: 5′-ACCCCAGAAGGCTACATTTTGC
CD44 exon 16: 3′-CTCATAGGACCAGAAGTTGTGG

Inclusion of individual alternative exons was monitored using 3′ primers specific for individual alternative exons and a primer complementary to constitutive exon 5:

CD44v5: (MASV5) 3′-TTGTAGCATGTGGGGTCTCCTCTT,
CD44v6: (MASV6) 3′-CCTTCTGTCACATGGGAGTCTTCA,
CD44v8: (MASV8) 3′-CACTGAAAGTGGTC CTGTCCTGTT.

All amplification reactions used an ATP-labeled 5′ primer. The PCR conditions for all amplifications were as follows: 25–35 cycles (see Figure legends) of 94°C for 1 min, 59°C for 1 min and 72°C for 1.5 min. Gel electrophoresis of 35 μl of the 100 μl PCR reaction was performed in 6% denaturing acrylamide gels at 34–40 mA. Expected amplification products are as follows: 142 bp (SRp20), 155 bp (hnRNP A1 without alternative exon 7B), 134 and 170 bp (hnRNP A2 with or without the alternative exon), and 221 bp (CD44 standard lacking any alternative exon). The expected amplification bands for alternatively-spliced CD44 mRNA containing only one variable exon are 207 bp (v5), 209 bp (v6) and 190 bp (v8). pBR 322/*Hpa*II markers were used in all displayed gels for product size determination. Identified amplification products resulting from the inclusion of one or two variable CD44 exons were sequenced to verify identity. After cloning of CD44 PCR products into a pCR 2.1 vector (TA cloning kit, Invitrogen, CA, USA) sequencing of these clones was performed using the thermo sequenase method (Amersham, OH, USA) and CD44 exon specific primers mentioned above.

Protein isolation and Western blot

Total cellular protein was isolated from the interphase and phenol phase from the initial homogenate after precipitation of the DNA with 0.3 ml 100% of ethanol per 1 ml of TRIzol used for the initial homogenization. The samples were stored at 26°C for 5 min and afterwards centrifuged at 2000 *g* for 5 min at 4°C. The phenol-ethanol supernatant was then precipitated with 1.5 ml isopropyl alcohol per 1 ml TRIzol used for initial homogenization. After storage for 10 min at 26°C the protein precipitates were centrifuged at 12 000 *g* at 4°C for 10 min. Protein pellets were washed three times in 0.3 M guanidine hydrochloride in 95% ethanol for 20 min at 26°C and centrifuged at 7500 *g* for 5 min at 4°C. The protein pellets were vortexed after the final wash in 2 ml 100% ethanol, stored for 20 min at 26°C and finally centrifuged at 7500 *g* for 5 min at 4°C. The pellets were resuspended in 1% SDS solution and incubated at 50°C for complete dissolution. Insoluble material was removed by centrifugation at 10 000 *g* for 10 min at 4°C. Supernatants were stored for further analysis at −80°C. Gel electrophoresis of 20 μg of total protein was performed using a 10%

SDS–PAGE gel at 100 volts. Afterwards the gels were electroblotted on a PVDF transfer membrane (PolyScreen, NEN Life Science, Boston, MA, USA) at 100 volts for 2.5 h at 4°C. After blocking in 5% Blotto/PBST membranes were incubated with mouse IgM mAb104 (1:5 dilution), the SRp20 specific mouse IgG Ab (1:200 dilution) or polyclonal rabbit anti-U2AF Ab 1:6000 dilution). The Western blots were stained by chemiluminescence (NEN Life Science) using appropriate horseradish-peroxidase labeled anti IgM antibody for mAb104 (Pierce, Rockford, IL, USA) or anti IgG antibody for SRp20 Ab and U2AF Ab a dilution of 1:3000.

Acknowledgements

This work was supported by Grant Sti 153/1-1 from the German Research Foundation (to E Stickeler), Grant DAMD 17-96-1-0684 from the US Army (to SM Berget) and Grant NCI CA 47112 from the National Cancer Institute (to D Medina).

References

Bell MV, Cowper AE, Lefranc MP, Bell JI and Screaton GR. (1998). *Mol. Cell. Biol.*, **18,** 5930–5941.

Berget SM. (1995). *J. Biol. Chem.*, **270,** 2411–2414.

Black DL. (1995). *RNA*, **1,** 763–771.

Cáceres JF, Stamm S, Helfman DM and Krainer AR. (1994). *Science*, **265,** 1706–1709.

Cáceres JF and Krainer AR. (1997). In: *Mammalian pre-mRNA splicing factors*. Krainer AR (ed.). IRL Press: Oxford, pp. 174–212.

Cannistra S, Abu-Jawdeh G and Niloff J. (1995). *J. Clin. Oncol.*, **13,** 1912–1921.

Cardiff RD. (1988). *Anticancer Res.*, **8,** 925–933.

Chandler SD, Mayeda A, Yeakley JM, Krainer AR and Fu X-D. (1997). *Proc. Natl. Acad. Sci. USA*, **94,** 3596–3601.

East JA and Hart IR. (1993). *Eur. J. Cancer*, **29A,** 1921–1922.

Fox SB, Fawcett J, Jackson DG, Collins I, Gatter KC, Harris AL, Gearing A and Simmons D. (1994). *Cancer Res.*, **54,** 4539–4546.

Friedrichs K, Kugler G, Franke F, Terpe HJ, Arlt J, Regidor PA and Gunthert U. (1995). *Lancet*, **345,** 1237.

Fu X-D. (1993). *Nature*, **365,** 82–85.

Fu X-D. (1995). *RNA*, **1,** 663–680.

Ge H and Manley JL. (1990). *Cell*, **62,** 25–34.

Ge H, Zuo P and Manley JL. (1991). *Cell*, **66,** 373–382.

Gunthert U, Hoffmann M, Rudy W, Reber S, Zoller M, Haussmann I, Matzku S, Wenzel A, Ponta H and Herrlich P. (1991). *Cell*, **65,** 13–24.

Gunthert U. (1993). *Curr. Microbiol. and Immunol.*, **184,** 47–63.

Haynes BF, Hua-Xin L and Patton KL. (1990). *Cancer Cells*, **3,** 347–350.

Heider KH, Mulder JW, Ostermann E, Susanni S, Patzelt E, Pals ST and Adolf GR. (1995). *Eur. J. Cancer*, **31A,** 2385–2391.

Jerry DJ, Ozbun MA, Kittrell FS, Lane DP, Medina D and Butel JS. (1993). *Cancer Res.*, **53,** 3374–3381.

Joensuu H, Klemi P, Toikkanen S and Jalkanen S. (1993). *Am. J. Path.*, **143,** 867–874.

Kittrell FS, Oborn CJ and Medina D. (1992). *Cancer Res.*, **52,** 1924–1932.

Krainer AR, Conway GC and Kozak D. (1990a). *Genes Dev.*, **4,** 1158–1171.

Krainer AR, Conway GC and Kozak D. (1990b). *Cell*, **62,** 35–42.

Kramer A. (1996). *Annu. Rev. Biochem.*, **65,** 367–409.

Kaufmann M, Heider KH, Sinn HP, von Minckwitz G, Ponta H and Herrlich P. (1995). *Lancet*, **345,** 615–619.

König H, Ponta H and Herrlich P. (1998). *EMBO J.*, **17,** 2904–2913.

Lee MP and Feinberg AP. (1997). *Cancer Res.*, **57,** 3131–3134.

Liu H-X, Zhang M and Krainer AR. (1998). *Genes Dev.*, **12,** 1990–2012.

Mackay C, Terpe HJ, Stauder R, Marston WL, Stark H and Gunthert U. (1994). *J. Cell. Biol.*, **124,** 71–82.

Manley J and Tacke R. (1996). *Genes Dev.*, **10,** 1569–1579.

Medina D. (1996). In: *Mammary Tumor Cell Cycle, Differentiation and Metastases*. Dickson RB and Lippman ME. (eds). Academic Publishers: Norwell, pp. 37–69.

Moore MJ, Query CC and Sharp PA. (1993). In: *The RNA World*. Gesteland RF and Atkins JF. (eds). Cold Spring Harbor Laboratory Press: New York, pp. 303–357.

Muller W, Schneiders A, Heider KH, Meier S, Hommel G and Gabbert HE. (1997). *J. Pathol.*, **183,** 222–227.

Neugebauer KM and Roth M. (1997). *Genes Dev.*, **11,** 1148–1159.

Norton PA. (1994). *J. Cell. Sci.*, **107,** 1–7.

Reed R. (1996). *Curr. Opin. Genet. Dev.*, **6,** 215–220.

Rio DC. (1993). *Curr. Opin. Genet. Dev.*, **3,** 574–584.

Screaton GR, Bell MV, Jackson DG, Cornelis FB, Gerth U and Bell JI. (1992). *Proc. Natl. Acad. Sci. USA*, **89,** 12160–12164.

Screaton GR, Cáceres JF, Mayeda A, Bell MV, Plebanski M, Jackson DG, Bell JI and Krainer AR. (1995). *EMBO J.*, **14,** 4336–4349.

Silberstein GB, Van Horn K, Strickland P, Roberts Jr CT and Daniel CW. (1997). *Proc. Natl. Acad. Sci. USA*, **94,** 8132–8137.

Sinn HP, Heider KH, Skroch-Angel P, von Minckwitz G, Kaufmann M, Herrlich P and Ponta P. (1995). *Breast Cancer Res. Treat.* **36,** 307–313.

Stamenkovic I, Amiot M, Pesando JM and Seed B. (1989). *Cell*, **56,** 1057–1062.

Stickeler E, Runnebaum IB, Moebus VJ, Kieback DG and Kreienberg R. (1997). *Anticancer Res.*, **17,** 1871–1876.

Sun Q, Mayeda A, Hampson RK, Krainer AR and Rottman FM. (1993). *Genes Dev.*, **7,** 2598–2608.

Tran TA, Kallakury BV, Sheehan CE and Ross JS. (1997). *Hum. Pathol.*, **28,** 809–814.

Valcarcel J and Green MR. (1996). *Trends Biochem. Sci.*, **21,** 296–301.

Wang J and Manley JL. (1995). *RNA*, **1,** 335–346.

Wielenga VJ, Heider KH, Offerhaus GJ, Adolf GR, van den Berg FM, Ponta H, Herrlick P and Pals ST. (1993). *Cancer Res.*, **53,** 4754–4756.

Zamore PD, Patton JG and Green MR. (1992). *Nature*, **355,** 609–614.

Zahler AM, Lane WS, Stolk JA and Roth MB. (1992). *Genes Dev.*, **6,** 837–847.

Zahler AM, Neugebauer KM, Lane WS and Roth MB. (1993a). *Science*, **260,** 219–222.

Zahler AM, Neugebauer KM, Stolk JA and Roth MB. (1993b). *Mol. Cell. Biol.*, **13,** 4023–4028.

Zhu X, Daffada AA, Chan CM and Dowsett M. (1997). *Int. J. Cancer*, **72,** 574–580.